Essays in Empirical Macroeconomics with Application to Monetary Policy in a Data-Rich Environment

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

Dr. rer. pol.
im Fach Volkswirtschaftslehre

eingereicht an der
Wirtschaftswissenschaftlichen Fakultät
Humboldt-Universität zu Berlin

von
Dipl.-Vw. Pooyan Amir Ahmadi
1. September 1980

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Christoph Markschies

Dekan der Wirtschaftswissenschaftlichen Fakultät:
Prof. Oliver Günther, Ph.D.

Gutachter:
1. Prof. Harald Uhlig, Ph.D.
2. Prof. Bartosz Maćkowiak, Ph.D.

eingereicht am: 23. April 2008
Tag der mündlichen Prüfung: 9. Juli 2009

To Friederike...

Acknowledgments

This thesis was written while I was working as a member of the Collaborative Research Center 649 Economic Risk at Humboldt Universität zu Berlin. I have been supported by many colleagues during the last two and a half years to whom I am grateful for supporting me.

First and foremost I would like to thanks my supervisors Professor Harald Uhlig, Professor Bartosz Mackowiak and Professor Albrecht Ritschl for continuous support, guidance and mentorship. This thesis would not have come to exist without their support.

I would like to thank my thesis supervisor Professor Harald Uhlig to whom I am very grateful for teaching me, providing valuable support, encouragement and comments. He significantly contributed to my thesis also directly as a co-author in Chapter 2 *Measuring the Dynamic Effects of Monetary Policy Shocks: A Bayesian FAVAR Approach with Sign Restriction*. Furthermore he provided numerous valuable suggestions and comments to all other chapters.

I would also like to express my gratitude to Professor Bartosz Maćkowiak who has been also supervising this thesis. I'm grateful to him for constant encouragement and teaching me Bayesian time series econometrics which was the starting point of my research interest in this field. During the last three years I have benefited enormously discussing all kinds of different questions with him and working on a joint project.

Professor Albrecht Ritschl has been a major source of support already supporting me as a student. He has always been a source of inspiration for conducting research. I would like to express my gratitude to him for endless mentorship, support and encouragement. I benefited from discussing all kinds of research questions and working with him in a joint project resulting in chapter 3 *Monetary Policy during the Great Depression*. He taught me a lot. Not only was I lucky to get academic support and advice he has also always been exerted to provide funding and encouraging me to attend conferences early on which was extremely helpful.

Parts of this thesis were written during my stay at Princeton University. I would like to particularly thank Professor Chris Sims and Professor Mark Watson for inviting me twice providing me with countless suggestions and comments. My thesis was shaped and improved substantially during the years that I was able to spend there. Moreover, I am highly indebted to Professor Harald Uhlig who strongly supported my research stay.

I also benefited from comments during seminar and conference presentations in Berlin, Princeton, Lund, Exeter, Budapest, Zurich and Prague. In particular Chapter 4 of this thesis was greatly improved by suggestions from Professor Chris Sims, Professor Mark Watson, Professor Nobuhiro Kiyotaki and seminar participants of the student macro seminar at Princeton University.

I am also grateful to Samad Sarferaz for endless discussions over the years who shares with me many research interests and therefore has always been an excellent scholar to discuss all kinds of questions. I would also like to thank Henning Weber who has been a great office mate and I was lucky to share my office with and discuss all kinds of thoughts and questions. Furthermore I would like to thank Holger Gerhardt, Martin Kliem and Stefan Ried for helping me with all kinds of organizational issues.

I thank the Deutsche Forschungsgemeinschaft for funding this research through the Collaborative Research Center (CRC) 649. Furthermore I would like to thank the DAAD (German Academic Exchange Service) for a Doctoral Scholarship. Maria Grith has provided excellent research assistance in the course of time.

Last but not least I would like to thank my family and friends who have been a constant source of support. I want to express my special gratitude to Friederike Westphal for being an endless and reliable source of support for over eleven years.

Abstract

This thesis consists of four self-contained chapters. The first chapter provides an introduction with a literature overview.
In Chapter 2 we estimate the effects of monetary policy shocks in a Bayesian Factor-Augmented vector autoregression (BFAVAR). We propose to employ as an identification strategy sign restrictions on the impulse response function of pertinent variables according to conventional wisdom. The key strength of our factor based approach is that sign restrictions can be imposed on many variables in order to pin down the impact of monetary policy shocks. Thus an exact identification of shocks can be approximated and monitored.
In chapter 3 the role of monetary policy during the interwar Great Depression is analyzed. The prominent role of monetary policy in the U.S. interwar depression has been conventional wisdom since Friedman and Schwartz [15]. This paper attempts to capture the pertinent dynamics through a BFAVAR methodology of the previous chapter. We find the effects of monetary policy shocks and the systematic component to have been moderate. Our results caution against a predominantly monetary interpretation of the Great Depression.
This final chapter 4 analyzes macroeconomic dynamics within the Euro area. To tackle the questions at hand I propose a novel approach to jointly estimate a factor-based DSGE model and a structural dynamic factor model that simultaneously captures the rich interrelations in a parsimonious way and explicitly involves economic theory in the estimation procedure. To identify shocks I employ both sign restrictions derived from the estimated DSGE model and the implied restrictions from the DSGE model rotation. I find a high degree of comovement across the member countries, homogeneity in the monetary transmission mechanism and heterogeneity in transmission of technology shocks. The suggested approach results in a factor generalization of the DSGE-VAR methodology of Del Negro and Schorfheide [12].

Zusammenfassung

Diese Dissertation besteht aus vier eigenständigen Aufsätzen. Das erste Kapitel liefert eine Einleitung uns einen Literaturüberblick.
Im zweiten Kapitel schätzen wir die Effekte eines geldpolitischen Schocks in einer Bayesianischen faktorerweiterten Vektorautoregression. Als ein Identifikationsschema schlagen wir theoretisch fundierte Vorzeichenrestriktionen vor, welche auf die angemessenen Impuls-Antwortfolgen auferlegt werden können. Der Vorteil der faktorbasierten Vorzeichenrestriktion liegt in der Möglichkeit sehr viele theoretische fundierte Restriktionen zu setzen um so exakter zu identifizieren.
Im dritten Kapitel untersuchen wir die Rolle der Geldpolitik während der Weltwirtschaftskrise in den USA. Die besondere Rolle der Geldpolitik gilt seit Friedman and Schwartz [15] als gängige Meinung. In diesem Papier versuchen wir die entscheidenden Dynamiken der Zwischenkriegszeit mit dem BFAVAR Modell abzubilden und die Effekte geldpolitischer Schocks zu analysieren. Weiterhin schauen wir uns die Effekte der systematischen Komponente der Geldpolitik an. Wir finden heraus, dass der Anteil der Geldpolitik insgesamt zwar präsent allerdings recht gemäßigt vorhanden.
Im vierten Kapitel werden die makroökonomischen Dynamiken innerhalb des Euroraumes untersucht. Hierbei schlage ich einen neuen Ansatz vor um die vielen relevanten Interrelationen effizient und sparsam zu vereinbaren. Ein faktorbasiertes DSGE Modell wird gemeinsam mit einem dynamischen Faktormodell geschätzt. Hierbei wird explizit ökonomische Theorie zur Datenanalyse verwendet. Zur Identifikation makroökonomischer Schocks verwende ich sowohl Vorzeichenrestriktionen wie auch die DSGE Rotation.

Contents

List of Figures

List of Tables

1 Literature

1.1 Objective of Study

This thesis provides contribution to the field of empirical macroeconomics in which a plethora of data is necessary, available and monitored by economic agents and policy makers prior to making decisions. Reflecting the information set in a realistic, flexible and parsimonious way has been the objective of recent advances and research efforts in the field in order to avoid biased inference only relying on few data. This is important e.g. to understand how monetary and fiscal policy affect the economy. The Traditional models like the vector autoregressions (VAR) are limited to cope with the large dimension of the data to be incorporated in the estimation procedure. Efforts to combine economic theory with the estimation and identification of large dimensional empirical models are promising, challenging and important to match theory and data.

This study focuses mainly on three questions that are challenging within a framework to cope with a large set of data. First, what are the effects of monetary policy shocks in the data-rich environment of the US post-World War II period? How sensitive, robust and reasonable are the results under a traditional and rather widespread factor generalized Cholesky identification compared to factor generalization of sign restriction to identify monetary policy shocks and in particular how do results change taking into account the different US monetary policy regimes at play. The second question analyzes the role and contribution of monetary factors in the US Great Depression by putting the Friedman and Schwartz hypothesis into an empirical test who assign a predominant role to monetary policy in the US Great Depression. The analysis involves a unique complete data set capturing the rich interrelations of the financial, monetary and real sector of the economy. The third question addresses the degree of comovement of the constituent countries of the Euro area and the single countries exposure to common macroeconomic shocks by characterizing the degree of heterogeneity of the transmission mechanism of shocks across the countries. To answer the questions raised we employ sophisticated methods of Bayesian time series econometrics, dynamic factor models, Markov chain Monte Carlo (MCMC) simulation methods and numerical methods each designed to cope with a rich set of informative data in a flexible and parsimonious way.

1.2 Outline of Study

My thesis mainly consists of three chapters answering different questions of distinct interest to answer interesting questions of the effects of macroeconomic policy on the macroeconomy. In chapter 2 we propose to measure the effect of monetary policy by employing sign restrictions on the impulse response functions of selected variables consistent with economic theory. We find that our proposed identification approach for dynamic factor models is promising producing robust and reasonable results and appealing even reliable under parameter instability. The traditional identification scheme like the recursive Cholesky approach are particularly sensitive to the choice of variable producing unreasonable result in particular when parameter instability is taken in to account. In chapter three the role of monetary policy during the US Great Depression is revisited analyzing a unique large panel of data. We find that systematic monetary policy was restrictive in late 1929 and again in 1931, however the effects were again quite moderate. Our results caution against a predominantly monetary interpretation of the Great Depression. Finally in Chapter four I suggest to combine to match theory to data by jointly estimating a DSGE and a DFM to measure the degree of comovement across the euro area countries and it's heterogeneity in the transmission mechanism of shocks. I find that the approach employing a large data set fits better compared to single indicators. Furthermore invoking the joint estimation of DSGE estimation improves the fit. I find a high degree if comovement and by and large homogeneity in the monetary transmission mechanism. The following paragraphs summaries in succinctly and densely the key findings. Details about estimation procedure can be found in the respective chapters' appendix.

To address the first question we employ the Bayesian factor-augmented VAR (FAVAR) framework strongly related to the literature on dynamic factor models analyzing a large panel of post-World War II US data. Instead of relying on few indicators as in the traditional VAR framework we employ the Bayesian dynamic factor analysis which can reflect realistically the information set central bankers have to base their policy decisions on. The importance arises as single indicators reflecting economic concepts may easily lead to biased inference, results are sensitive to the choice of these indicators, missing data and dynamics lead to omitted variable bias and the few indicators can only give a limited picture of the effects of macroeconomic shocks on the macroeconomy. The enticing promise of the chosen class of models are the inherent broader view of the effect of the economy on a disaggregate level. To identify the shocks we employ a generalized recursive Cholesky identification as it has been done mostly in the literature before. We provide an alternative relying on sign restrictions on the impulse response functions of selected variables. One key advantage and promising feature of this approach is that in the large dimensional framework one can impose a larger set of indicators that are soundly chosen according to conventional wisdom satisfying economic theory agreed on. Furthermore the two approaches disadvantages, robustness and advantages are tested under parameter instability by analyzing the data for the different monetary policy regimes at play. The importance and sensitivity of identification is addressed. The question at hand is left as open as possible to avoid circular reasoning. We find that sign restriction approach is robust and a sensible choice also under parameter instability. We show that the recursive Cholesky identification produces strong and persistent price puzzles for the post-Volcker disinflation period. Further we find some small real effect of contractionary monetary policy shocks showing a slight decrease in some output indicators though their contribution in output fluctuations are modest.

For analyzing the second question about the monetary factors during the US Great Depression we compile a unique large and complete panel of the US interwar data covering the real, financial and monetary sector on a disaggregate level. We employ the Bayesian FAVAR model combined with the proposed sign restriction approach for the non-systematic component of monetary policy. We invoke Granger causality of monetary instruments, implemented by Bayesian forecasting techniques, to identify the systematic component of monetary policy. We attempt to unravel, quantify and evaluate the role of monetary policy and its contribution to the US Great Depression. We broadly confirm the Friedman and Schwartz view that systematic monetary policy was restrictive in late 1929 and again in 1931. However, the effects were again quite moderate. Our results caution against a predominantly monetary interpretation of the Great Depression.

To answer the third question about the degree of comovement and the heterogeneity of the euro are transmission mechanism of shocks I propose a new method based on recent advances. It involves the joint estimation of a dynamic stochastic general equilibrium model (DSGE) and a dynamic factor model (DFM) resulting in an new class of models in spirit of the DSGE-VAR literature (see Del Negro and Schorfheide [12]). The approach explicitly invokes economic theory into the estimation and identification. The attempt is to quantify how much the euro are countries are driven by common shocks. Furthermore I want to shed light on the exposure of the single countries to the common aggregate structural shocks. For the estimation we employ Markov Chain Monte Carlo methods in particular a Metropolis-within-Gibbs sampler. To identify the macroeconomic shocks and evaluate the degree of heterogeneity I employ robust model-based sign restrictions and propose a new method to combine DSGE and DFM estimation resulting into a new class of models namely the DFM-DSGE framework. I find a high degree of comovement among the countries suggesting that single member countries are largely driven by common shocks. Furthermore I find by and large a homogeneity in the monetary transmission mechanism though there are some differences in the labor market of some countries. Overall based on Bayesian model choice I find first that the fit of models improve if a large panel of the single countries pertinent macroeconomic variables is included opposed to the euro area aggregates. Second the new proposed DSGE-DFM model fits better than the pure DFM counterpart.

2 Measuring the Dynamic Effects of Monetary Policy Shocks: A Bayesian FAVAR Approach with Sign Restriction

Joint with Harald Uhlig

In this paper we estimate the effects of monetary policy shocks in a Bayesian Factor-Augmented vector autoregression (BFAVAR). We propose to employ as an identification strategy sign restrictions on the impulse response function of pertinent variables according to conventional wisdom, namely prices, monetary aggregates, spreads and interest rates. The key strength of our factor based approach is that sign restrictions can be imposed on many variables in order to pin down the impact of monetary policy shocks. Thus an exact identification of shocks can be approximated and monitored. We find that our factor generalization of sign restriction outperforms the competing Cholesky identification, is robust across different subsamples and avoids anomalies present in Cholesky identification such as a unreasonable "prize puzzles" by construction.

2.1 Introduction

What are the dynamic effects of monetary policy shocks throughout the economy? In this paper we answer the question at hand by combining two recent advances in empirical macroeconomics, namely factor-augmented VARs (FAVAR) with sign restrictions to measure the effects of monetary policy shocks. Identification schemes designed to cope with this class of large dimensional models including unobserved components are rare mostly relying on recursive Cholesky identification schemes with the policy instrument ordered last. We show in different setups that this approach fails to correctly identify monetary policy shocks leading to flawed unreasonable results particularly for the post-Volcker disinflation period. Our approach is promising as it offers to select from a large set of variables to be imposed though soundly chosen subject to economic theory. Furthermore robustness can be monitored easily by comparing results for many indicators related to the same economic concept.

Though prevalently analyzed in a large body of literature the effects of monetary policy shocks on the real side of the economy is still subject to debate. A major suspect for the dissent regarding results and conclusion is the key ingredient of identification. Depending on the assumptions underlying the identifying strategy, results are different sometimes at odds with economic theory and the class of (log-)linearized DSGE models. To grapple the task at hand we propose to employ Bayesian Factor-Augmented vector autoregression (henceforth BFAVAR) identified via a richer set of sign restrictions than currently employed in the literature. We argue that our approach is well suited to successfully pin down the correct impact of a large number of unconstrained variables of interest. To this end we can approximate an exact identification by setting restrictions grounded on widely accepted conventional wisdom derived from economic theory. This amounts to impose a negative response of prices, money and a positive response for short term interest rates for a specified period following a contractionary monetary policy shock.

Following the lead of Sims (Sims [25]) most researchers analyze the question at hand through the lens of a vector autoregression (VAR). Most VAR studies consider a small number of variables in order to save degrees of freedom for keeping the model tractable. This means with few exceptions to employ a 6-8 variable VAR[1]. As pointed by Bernanke and Boivin [4] monetary policy takes place in a "data-rich environment" a feature that VARs cannot accomplish appropriately due to the "curse of dimensionality" they suffer from. Thus the appealing feature of tractability also marks the limitations inherent in (monetary) VARs of small scale.

There are four key limitations of conventional VARs which dynamic factor models (henceforth DFM) can cope with, hence motivating our choice for the FAVAR framework. First, the restricted set of variables considered in VAR models is at odds with the rich set of information available to and obviously monitored by the private sector and central bankers prior to taking their decisions. Thus this restriction can entail the well known "omitted variable bias" which might lead to biased inference. Anomalies such as the "price puzzle", raised by Sims [26] is argued to be indebted to the lack of inappropriate controlling for information that central bankers probably have about future inflation. Second, those few indicators considered in the VAR analysis are each supposed to represent a whole economic concept, e.g. GDP is supposed to reflect economic activity which is apparently restrictive. This reflects a fragmentary picture of the dynamics of the economy and hence is inept. Third, we are restricted to check the impact of only those few variables considered. However, academic researchers and practioneers alike are interested and concerned about many more economic variables affected by macroeconomic shocks. Forth, and of most interest for our identification strategy, is that the larger number of time series considered paves the way for exact identification as many more indicators can be restricted according to economic theory. Hence we can pin down the effects of a random shift in monetary policy with less uncertainty as the identified responses fulfill a broader set of economically reasonable assumptions. For instance many more price series could be restricted not to increase after a contractionary monetary policy shock. Thus robustness of identification can be monitored.

Recent advances in DFM help to overcome the aforementioned drawbacks by parsimoniously extracting few dynamic factors out of a large panel of macroeconomic data, which summarize the crucial comovements among the driving forces of the economy. Thus the rich interrelation of the monetary, financial and the real sector are detected. Bernanke and Boivin [4] and Bernanke, Boivin, and Eliasz [6] coined the FAVAR models which is a unifying framework that combines DFM with the VAR analysis. This approach has been surveyed and extended with respect to different identification strategies for the classical approach by Stock and Watson [28]. They argue that the large set of data considered both reflects the monetary transmission mechanism better and improves the identification of shocks. For the estimation we choose the Bayesian

[1] Leeper, Sims, and Zha [22] employ a 18 Bayesian VAR imposing over identifying restrictions.

likelihood-based estimation based on MCMC methods which is fully parametric. Thus we can explicitly exploit the factor structure of the data and the law of motion of the extracted factors. This comes at the cost of computational burden.

The key ingredient for analyzing the impact of a random shift in monetary policy is identification which is a highly controversial matter. Different assumptions produce different results sometimes at odds with economic theory. We promote to rather explicitly impose economic theory instead of implicitly expecting via the tool of sign restrictions. Key advantages are the following. First, sign restrictions, as introduced by Uhlig [31] impose "conventional wisdom" explicitly as part of the identifying assumptions. Second, sign restrictions avoid anomalies such as the "price puzzle" by construction. Third, for identification purposes DSGE models contain a large number of adaptive sign restrictions while they seldom deliver the whole set of zero restrictions required to recover all economic shocks. Hence it is a "weak" identification compared to schemes based on zero-type restrictions. Finally, it is quite easy to implement sign restrictions ex post.

As a comparison for the performance of our approach we choose the Cholesky identification applied by Bernanke, Boivin, and Eliasz [6]. They prefer the results provided by the nonparametric two-step estimation approach based on dynamic principal components. Amongst others they argue the results seem more reasonable to them. We show that the problem with their likelihood-based results is that the employed Cholesky identification does not identify correctly the structural shock. In the literature is a tendency of uniformity w.r.t. the underlying assumptions in applied work towards the Cholesky identification. But this approach is unrelated to the class of DSGE models and hence results are often at odds with what economic theory predicts. Cholesky identification relies on the informational orderings about the arrival of shocks. For a discussion of the criticism of Cholesky identification see, e.g. Cooley and LeRoy [1985] who argue that the contemporaneous recursive structure is hard to obtain in general equilibrium. Canova and Pina [2005] show that DSGE models almost never recover the zero restrictions employed to identify monetary shocks and that misspecification of the feature of the model economy can be substantial.

Identifying via Cholesky identification with the policy instrument ordered last we find that the price puzzle is still present and unreasonably high and in our benchmark model even increasing for a period up to four years. We conclude that the Cholesky identification does not recover the structural shocks. Hence the results are not identified. We find that industrial production decreases after a monetary contraction with a maximum impact after two years before reverting back. However this result is sensible with respect to the data span considered for estimation. Our subsample analysis shows that this result vanishes for the post Volcker disinflation period. Hence for the latter case the impact on output is not inconsistent with monetary neutrality. Regarding the FAVAR specification we find that our benchmark model which includes the federal funds rate and CPI as factors included in the FAVAR equation combined with our "maximal set of restrictions" works best as regards the uncertainty associated[2]. Furthermore we find that the forecast-error revision variance of output due to random shifts in

[2]It is important to note that this benchmark specification is different to the one of BBE who only include the federal funds rate as a factor in the FAVAR equation. More details about our choice can be found on the section on FAVAR specification.

monetary policy accounts for less than 20 percent on average which is consistent with results in Sims and Zha [27] and Uhlig [31]. Similar results have been obtained for the period of the US Great Depression in Sims [24] and Amir Ahmadi and Ritschl [1].

There different approaches to the sign restriction identification which are all common in the sense that they do not rely on zero-type restriction and identify via restriction on the sign and/or shape of the structural impulse response function. But they are different in their motivation. Early references are Faust [14], Canova and Pina [9], Canova and Nicolo [8]. Faust [14] imposes sign restrictions only on impact focusing on the fragility of the consensus conclusion that monetary policy shocks account for a small fraction of output fluctuations. Canova and Nicolo [8] and Canova and Pina [9] impose restrictions on the cross-correlation of variables in response to shocks, adding restrictions until the maximum number of shocks is uniquely identified. Uhlig [31] imposes the restriction for several periods and leaves the variable of key interest in the analysis unrestricted, hence the term agnostic identification.

2.2 The Model

The model we apply is the Bayesian version of the FAVAR model introduced by Bernanke, Boivin, and Eliasz [6] which is cast in state space form. Here X_t^c denotes the $[N_c \times 1]$ vector of observable variables in period t, where $t = 1, \ldots, T$ is the time index and the superscript c refers to the panel out of which the common factors are extracted. Let f_t^c denote the $[K_c \times 1]$ vector of unobservable factors in period t and e_t^c the $[N_c \times 1]$ time t idiosyncratic component of the respective variables. Furthermore let f_t^y denote the $[N_y \times 1]$ perfectly observable vector of variables that have pervasive effects throughout the economy and are considered so important that they should be included as factors. N_c, K_c and N_y denote the number of variables in X_t^c, the number of factors to be extracted from X_t^c and the number of perfectly observable factors respectively. The model is

$$X_t^c = \lambda^c f_t^c + \lambda^y f_t^y + e_t^c \tag{2.2.1}$$

$$e_t^c \sim N(0, R_e) \tag{2.2.2}$$

Here λ^c and λ^y denote the matrix of factor loadings of the factors and the perfectly observable variables included as factors with dimension $[N_c \times K_c]$ and $[N_c \times N_y]$ respectively. The error term e_t^c has mean 0 and covariance R_e which is assumed to be diagonal. Hence the error terms of the observable variables are mutually uncorrelated. The FAVAR state equation represents the joint dynamics of factors and the observable policy variables (f_t^c, f_t^y) following a $VAR(P)$ process.

$$\begin{bmatrix} f_t^c \\ f_t^y \end{bmatrix} = \sum_{p=1}^{P} \phi_p \begin{bmatrix} f_{t-p}^c \\ f_{t-p}^y \end{bmatrix} + u_t^f \tag{2.2.3}$$

$$u_t^f \sim N(0, Q_u) \tag{2.2.4}$$

where u_t^f is the time t reduced form shock, Q_u is the factor error covariance matrix and the ϕ_p's denote the respective p-lag coefficient matrices. The dimensions are $[K \times 1]$,

$[K \times 1]$ and $[K \times K]$ respectively. Note that the total number of factors is $K = K_c + N_y$.

2.3 Estimation and Inference

For the estimation of FAVAR models Bernanke, Boivin, and Eliasz [6] present two competing approaches. The first one, which they prefer due to their results and the computational simplicity is the two-step estimation based on a dynamic principal component approach. This classical approach goes back to Stock and W.Watson [29], Stock and Watson [28]. A detailed survey with several identification schemes in the classical estimation approach can be found in Stock and Watson [28]. The second estimation approach described in their paper is the one that we employ in this paper because the likelihood based one-step estimation approach employing MCMC methods explicitly exploits the factor structure.

We pursue the multi-move Gibbs sampler for which we have to cast the model into the state space representation. Let (2.2.3) be extended by f_t^y which results in

$$\begin{bmatrix} X_t^c \\ f_t^y \end{bmatrix} = \begin{bmatrix} \lambda^c & \lambda^y \\ 0 & I_{N_y} \end{bmatrix} \begin{bmatrix} f_t^c \\ f_t^y \end{bmatrix} + \begin{bmatrix} e_t^c \\ 0 \end{bmatrix} \tag{2.3.1}$$

Let $X_t \equiv (X_t^{c\prime}, f_t^{y\prime})'$, $e_t \equiv (e_t^{c\prime}, 0_{[N_y \times 1]}{}')'$ and $f_t \equiv (f_t^{c\prime}, f_t^{y\prime})'$, then the model can be rewritten as

$$X_t = \lambda f_t + e_t \tag{2.3.2}$$

$$f_t = \sum_{p=1}^{P} \phi_p f_{t-p} + u_t^f \tag{2.3.3}$$

where X_t has dimension $[N \times 1]$ with $N = N_c + K_y$. In most empirical applications and also in our specification the lag order P exceeds one hence we have to rewrite the state space in a stacked first order Markov process. This requires the following straight forward definitions for the companion form of the model:

$$\begin{aligned} \lambda &\equiv \begin{bmatrix} \lambda^c & \lambda^y \\ 0_{N_y \times K_c} & I_{N_y} \end{bmatrix} \\ \phi &\equiv [\phi_1, \phi_2, \ldots, \phi_P]' \\ F_t &\equiv (f_t, f_{t-1}, ..., f_{t-p+1}) \\ u_t &\equiv (u_t^f, 0, ..., 0)' \end{aligned}$$

The lag polynomial of the FAVAR equation in the first-order representation changes to:

$$\Phi = \begin{bmatrix} \phi_1 & \cdots & \phi_P \\ I_{K(P-1)} & & 0_{K(P-1) \times K} \end{bmatrix}.$$

Now we have to transform the VCV of the FAVAR disturbances with 0's in a straightforward way to adjust the dimensions of the state equation which results in the following matrix:

$$Q = \begin{bmatrix} Q_u & 0 \\ 0 & 0 \end{bmatrix}$$

where Q is of dimension $[PK \times PK]$ extended by zero matrices to match the companion form. We define $\Lambda \equiv [\lambda \quad 0 \quad \cdots \quad 0]$. Then

$$F_t = \Phi F_{t-1} + u_t \tag{2.3.4}$$

$$X_t = \Lambda F_t + e_t \tag{2.3.5}$$

$$u_t \sim \mathcal{N}(0, Q) \tag{2.3.6}$$

$$e_t \sim \mathcal{N}(0, R) \tag{2.3.7}$$

is the final state-space representation prepared to fit the estimation procedure. Note again that R is diagonal and that e_t and u_t are mutually independent.

2.3.1 Factor Identification

The factors are only identified up to an invertible rotation. Any rotation of the factors results in the same likelihood for the factors though the models are different. Identifying restrictions have to be set, in order to distinguish the idiosyncratic from the common component. Additionally one can set further identifying assumptions in order to identify the factors and the loadings, separately. We follow the standard identification restrictions either on the factor loading matrix employed by Bernanke, Boivin, and Eliasz [6] for unique identification against rotational indeterminacy. Since factors are estimated up to a rotation, the normalization should not affect the space spanned by the estimated factors. In the joint estimation case the specified identification against rotation requires that the factors are uniquely identified in the following form

$$f_t^* = A f_t^c - B f_t^y \tag{2.3.8}$$

where A and B are nonsingular. Restrictions are only imposed on the observation equation. Here we substitute F_t^* into (2.2.1) due to the fact that restrictions should not be imposed on the VAR dynamics we obtain

$$X_t^c = \lambda^c A^{-1} f_t^* + (\lambda^y + \lambda^c A^{-1} B) f_t^y + e_t^c. \tag{2.3.9}$$

For unique identification of the factors and the loadings it is required that $\lambda^c A^{-1} = \lambda^c$ and $\lambda^y + \lambda^c A^{-1} B = \lambda^y$. As discussed in Bernanke, Boivin and Eliasz [2005] sufficient conditions are to set the upper $K_c \times K_c$ block of λ^c to identity and the upper $K_c \times N_y$ block of λ^y to a zero matrix[3].

2.3.2 Inference

Bayesian analysis treats the parameters of the model as random variables. We are interested in inference on the parameter space $\theta = \left(\lambda^f, \lambda^y, R_e, \phi, Q_u\right)$ and the factors

[3]Note that this identification strategy is over-identified. However, for comparison purposes we follow closely the approach of Bernanke, Boivin, and Eliasz [6].

$\{f_t\}_{t=1}^{T}$. Multi move Gibbs Sampling alternately samples the parameters θ and the factors f_t, given the data. We use the multi move version of the Gibbs sampler because this approach allows us, as a first step, to estimate the unobserved common components, namely the factors via the Kalman filtering technique conditional on the given hyperparameters and data, and as a second step calculate the hyperparameters of the model given the factors and data via the Gibbs sampler in the respective blocking. Let $X^T = (X_t, \ldots, X_T)$ and $F^T = (F_1, \ldots, F_T)$ define the respective histories. For the estimation of the model we want to derive the posterior densities which requires to empirically approximating the marginal posterior densities of F^T and θ:

$$p(F^T) = \int p(F^T, \theta) d\theta$$
$$p(\theta) = \int p(F^T, \theta) dF^T$$

where

$$p(F^T, \theta)$$

is the joint posterior density and the integrals are taken with respect to the supports of θ and F^T respectively. The procedure applied to obtain the empirical approximation of the posterior distribution is the previously mentioned multi move version of the Gibbs sampling technique by Carter and Kohn [10] and Frühwirth-Schnatter [16][4].

2.3.3 Choosing the Starting Values

In general one can start the iteration cycle with any arbitrary randomly drawn set of parameters, as the joint and marginal empirical distributions of the generated parameters will converge at an exponential rate to its joint and marginal target distributions as $S \to \infty$. This has been shown by Geman and Geman [18]. Since Gelman and Rubin [17] have shown that a single chain of the Gibbs sampler might give a "false sense of security ", it has become common practice to try out different starting values. We check our results based on four different strategies regarding the set of starting values. One out of many convergence diagnostics involves testing the fragility of the results with respect to the starting values. For the results to be reliable, estimates based on different stating values should not differ. Strictly speaking, the different chains should represent the same target distribution. In order to verify we start our Gibbs sampler with the following summarized starting values respectively.

(i) Randomly draw θ_0 from (over)dispersed distribution

(ii) Set θ_0 to rather "agnostic values" which involves setting 0's for coefficients and 1's for variances[5]

(iii) Set θ_0 to results from principal component analysis.[6]In such a way the number of draws required for convergence can be reduced considerably.

[4]For a survey and more details see Kim and Nelson [21], Eliasz [13] and Bernanke, Boivin, and Eliasz [6]

[5]This strategy has been applied by Belviso and Milani [3].

[6]This strategy is particularly suited for large models as the ones studied here and has been proposed by Eliasz [13].

(iv) Set θ_0 to parameters of the last iteration of the previous run.

Despite the strategies above convergence is never guaranteed, particularly in large models. Hence it is recommended to restart a chain many times applying the strategy 4.

2.3.4 Conditional density of the factors F^T given X^T and θ

In this subsection we want to sample from $p(F^T \mid X^T, \theta)$ assuming that the data and the hyperparameters of the parameter space θ are given, hence we describe Bayesian inference on the dynamic evolution of the factors f_t conditional on X_t^c for $t = 1, \ldots, T$ and conditional on θ. The transformations that are required to draw the factors have been done in the previous section. The conditional distribution, from which the state vector is generated, can be expressed as the product of conditional distributions by exploiting the Markov property of state space models in the following way

$$p(F^T \mid X^T, \theta) = p(F_T \mid X^T, \theta) \prod_{t=1}^{T-1} p(F_t \mid F_{t+1}, X^T, \theta)$$

The state space model is linear and Gaussian, hence we have:

$$F_T \mid X^T, \theta \sim N(F_{T|T}, P_{T|T}) \tag{2.3.10}$$
$$F_{t|T} \mid F_{t+1|T}, X^T, \theta \sim N(F_{t|t,F_{t+1|T}}, P_{t|t,F_{t+1|T}}) \tag{2.3.11}$$

with

$$F_{T|T} = E(F_T \mid X^T, \theta) \tag{2.3.12}$$
$$P_{T|T} = Cov(f_T \mid X^T, \theta) \tag{2.3.13}$$
$$F_{t|t,F_{t+1|T}} = E(F_t \mid F_{t|t}, F_{t+1}, \theta) \tag{2.3.14}$$
$$P_{t|t,F_{t+1|T}} = Cov(F_t \mid F_{t|t}, F_{t+1}, \theta). \tag{2.3.15}$$

We first run the Kalman filter generating $F_{t|t}$ and $P_{t|}$ for $t = 1, \ldots, T$. For the initialization we set $F_{1|0} = 0_{KP \times 1}$ and $P_{1|0} = I_{KP}$ and iterate through the Kalman filter according to

$$F_{t|t} = F_{t|t-1} + P_{t|t-1}\Lambda' H^{-1} \eta_{t|t-1} \tag{2.3.16}$$
$$P_{t|t} = P_{t|t-1} - P_{t|t-1}\Lambda' H^{-1} \Lambda P_{t|t-1} \tag{2.3.17}$$

where $\eta_{t|t-1} = (X_t - \Lambda F_{t|t-1})$ is the conditional forecast error and its covariance is denoted by $H_{t|t-1} = (\Lambda P_{t|t-1}\Lambda' + R_e)$. Furthermore let

$$F_{t|t-1} = \Phi F_{t-1|t-1} \tag{2.3.18}$$
$$P_{t|t-1} = \Phi P_{t-1|t-1}\Phi' + Q_u. \tag{2.3.19}$$

The last iteration of the Kalman filter yields $F_{T|T}$ and $P_{T|T}$ required for (2.3.10) to draw the last observation and start the Kalman smoother according to (2.3.11) going backwards through the sample for F_t, $t = T-2, T-3, \ldots, 1$ updating the filtered estimates with the sampled factors one period up subject to

$$F^*_{t|t,F_{t+1|T}} = F_{t|t} + P_{t|t}\Phi^{*\prime} J^{-1}_{t+1|t}\xi_{t+1|t} \tag{2.3.20}$$

$$P^*_{t|t,F_{t+1|T}} = P_{t|t} - P_{t|t}\Phi^{*\prime} J^{-1}_{t+1|t}\Phi^* P_{t|t}. \tag{2.3.21}$$

where $\xi_{t+1|t} = F^*_{t+1} - \Phi^* F_{t|t}$ and $J_{t+1|t} = \Phi^* P_{t|t}\Phi^* + Q^*$. Note that Q^* refers to the upper $K \times K$ block of Q and Φ^* and F^*_t denote the first K rows of Φ and F_t respectively. This is required when Q is singular which is the case for the companion form when there is more than one lag in (2.3.3). Here we closely follow Kim and Nelson [21] where a detailed explanation and derivation can be found.

2.3.5 Conditional density of the parameters θ given X^T and F^T

Sampling from the conditional distribution of the parameters $p(\theta \mid X^T, F^T)$ requires the blocking of the parameters into the two parts that refer to the observation equation and to the state equation respectively. The blocks can be sampled independently from each other conditional on the extracted factors and the data.

2.3.5.1 Conditional density of Λ and R_e

This part refers to observation equation of the state space model which, conditional on the estimated factors and the data, specifies the distribution of Λ and R_e. The errors of the observation equation are mutually orthogonal with diagonal R_e. Hence we can apply equation by equation OLS in order to obtain the ols estimates $\hat{\Lambda}_n$ and $\hat{e}^c$ as the observation equation amounts to a set of independent regressions. Note that the subscript n refers to the n-th equation and all hat variables refer to the respective ols estimates. We assume conjugate priors

$$\begin{aligned} p(R_{nn}) &= \mathcal{IG}(\delta_0/2, \eta_0/2) \\ p(\Lambda_n \mid R_{nn}) &= \mathcal{N}(\Lambda_{n0}, R_{nn}M^{-1}_{n0}) \end{aligned}$$

which according to Bayesian results conform to the following conditional posterior distribution

$$\begin{aligned} p(R_{nn} \mid \tilde{X}_T, \tilde{F}_T) &= \mathcal{IG}(\delta_i/2, \eta_i/2) \\ p(\bar{\Lambda}_{nn} \mid \tilde{X}_T, \tilde{F}_T, R_{nn}) &= \mathcal{N}(\bar{\Lambda}_n, R_{nn}M^{-1}_n). \end{aligned}$$

with

$$\begin{aligned} \eta_n &= \eta_0 + T \\ \delta_n &= \delta_0 + \hat{e^c}'_n \hat{e^c}_n + (\hat{\Lambda}_n - \Lambda_{n0})' \left[M^{-1}_{n0} + (F^{n\prime}_T F^n_T)^{-1} \right]^{-1} (\hat{\Lambda}_n - \Lambda_{n0}) \\ \bar{M}_n &= M_{n0} + \left(F^{n\prime}_T F^n_T \right) \\ \Lambda_n &= \bar{M}_n (M^{-1}_{n0}\Lambda_{n0} + \left(F^{n\prime}_T F^n_T \right) \hat{\Lambda}_n) \end{aligned}$$

where we set the same prior specification ($\delta_0 = 6, \eta_0 = 10^{-3}, M_{n0} = I_{K_c}, \Lambda_{n0} = 0_{K_c \times 1}$) as in Bernanke, Boivin, and Eliasz [6] in order to allow an adequate comparison. M_0

denotes the matrix in the prior on the coefficients of the n-th equation of Λ_n. The factor normalization discussed earlier requires to set $M_0 = I$. The regressors of the n-th equation are represented by F_T^n and the fitted errors of the n-th equation are represented by $\hat{e}^c_{t\,n}$.

2.3.5.2 Conditional density of $vec(\phi)$ and Q_u

The next Gibbs block requires to draw $vec(\phi)$ and Q_u conditional on the most current draws of the factors and the data. We employ the Normal-Inverse Wishart prior according to Uhlig [30]

$$\begin{aligned} p(Q_u) =& \mathcal{IW}(S_0, \nu_0) \\ p(vec(\phi) \mid Q_u) =& \mathcal{N}(\bar{\phi}_0, Q_u \otimes N_0^{-1}) \end{aligned}$$

which results in the following posterior:

$$\begin{aligned} p(Q_u \mid X^T, F^T) =& \mathcal{IW}(S_T, \nu_T) \\ P(vec(\phi) \mid X^T, F^T, Q_u) =& \mathcal{N}(vec(\bar{\phi}_T), Q_u \otimes N_T^{-1}) \end{aligned}$$

with

$$\begin{aligned} \nu_T =& T + \nu_0 \\ N_T =& N_0 + (F'_{T-1}F_{T-1}) \\ \bar{\phi}_T =& N_T^{-1}(N_0\bar{\phi}_0 + F'_{T-1}F_{T-1}\hat{\phi}) \\ S_T =& \frac{\nu_0}{\nu_T}S_0 + \frac{T}{\nu_T}\hat{Q}_u + \frac{1}{\nu_T}(\hat{\phi} - \bar{\phi}_0)'N_0(N_T)^{-1}(F'_{T-1}F_{T-1})(\hat{\phi} - \bar{\phi}_0) \end{aligned}$$

This prior and has the following specification

$$\begin{aligned} \nu_0 =& K + 2 \\ N_0 =& 0_{K \times K} \end{aligned}$$

where the choice of S_0 and $\bar{\phi}_0$ are arbitrary as they cancel out in the posterior. We alternatively also implemented the Normal-Wishart prior for according to Kadiyala and Karlsson [20] where the diagonal elements of Q_0 are set to the corresponding p-lag univariate autoregressions, σ_i^2. The diagonal elements of Ω_0 are constructed such that the prior variances of the parameter of the k lagged j'th variable in the i'th equation equals $\sigma_i^2/k\sigma_j^2$. Hence $S_0 = Q_0$ and $\bar{\phi}_0 = \mathbf{0}$. Results were virtually the same. To ensure stationarity, we truncate the draws by discarding the draws of ϕ with the larges eigenvalue greater than 1 in absolute value.

2.4 Identification

The major objective of this paper is to identify monetary policy shocks in a data rich environment through imposing sign restrictions as introduced by Uhlig [31] for the VAR framework. The issue of how to identify structural shocks through the decomposition of the prediction error u_t^c and in particular of monetary policy shocks, has been subject of much debate in the literature. From an economic perspective it seems desirable to have an identification scheme which guarantees that the impulse responses satisfy conventional wisdom. In FAVAR models the task is actually the same with the main distinction that the structural shocks are not required to be deduced from the innovation of the reduced form observation equation, but from the FAVAR innovation, including the factors that summarize the crucial dynamics of the observed data. For comparison purposes we will employ two identifying schemes, namely the factor generalization of the aforementioned sign restriction and the factor generalization of the recursive Cholesky identification. These two identification schemes shall be explained in the following.

2.4.1 Factor generalization of Sign Restrictions

Identification of structural shocks through imposing sign restrictions is based on the assumptions about the sign of the impulse response functions for a specified period of key macroeconomic variables. Such restrictions should represent "conventional wisdom" derived from economic theory that most researchers can agree on. Conventional wisdom says that after a monetary policy contraction the federal funds rate should increase, the prices should not increase and the nonborrowed reserves should decrease. In other identification schemes that do not accomplish this wisdom, the researchers tend to call this empirical observation a "puzzle". There are even some researchers that try to build a model producing such "puzzles" out of a model as has been done by Christiano, Eichenbaum, and Evans [11]. Sims gives the advice to avoid unreasonable identification schemes. The sign restriction approach seems reasonable, especially in the context of FAVAR that can account for large cross-sectional data. Such models allow to check the robustness of identification and tighten the restriction along the lines of conventional wisdom as there are far more relevant indicators of interest.

In the context of a contractionary monetary policy shocks Uhlig [31] set the restriction that prices and nonborrowed reserves should not increase and the federal funds rate should increase for a specified period following a contractionary monetary policy shock. We employ different versions for our empirical analysis which is summarized in table (2.1). The structural version of the FAVAR equation (2.3.3)

$$f_t = \sum_{p=1}^{P} \phi_p f_{t-p} + u_t^f$$

where the matrix A is an orthogonal invertible matrix of dimension $[K \times K]$ which satisfies $u_t^c = A v_t^c$. We are only interested in identifying one single shock therefore it is sufficient to identify one single row of A denoted by a_k where k refers to the respective shock or row in A. This single-equation identification is the more common approach that most of the recent literature pursue. The alternative would be to identify one row but the whole matrix A, which means to identify the full system. This approach goes back to Blanchard and Watson [7]. The crucial step is to represent the one-step ahead prediction error u_t^c as a linear combination of orthogonalized structural shocks. We assume the fundamental innovations to be mutually independent and normalized to have unit variance. Hence $E[v_t^c v_t^{c\prime}] = I_K$. The restriction on A emerges from the covariance structure of the reduced form factor innovation which results in

$$\begin{aligned} Q_u =& E[u_t^f u_t^{f'}] \\ E[u_t^f u_t^{f'}] =& AE[v_t^f v_t^{f'}]A' \\ AE[v_t^f v_t^{f'}]A' =& AA' \end{aligned}$$

According to Uhlig [31] we define an impulse vector as a column of the matrix A. Such a vector can be obtained from any decomposition, e.g. the Cholesky decomposition, of the covariance matrix of the factor residual matrix $\tilde{A}\tilde{A}' = Q_u$.

Definition 1 *The vector $a \in \Re^K$ is called an* ***i****mpulse vector, iff there is some matrix A, so that $AA' = Q_u$ and so that a is a column of A*

According to the Proposition 1 of Uhlig [2005,pp. 18], any impulse vector can be characterized as follows. Let $\tilde{A}\tilde{A}' = Q_u$ be the Cholesky decomposition. Then a is an impulse vector if and only if there is some K-dimensional vector α of unit length so that

$$a = \tilde{A}\alpha$$

Given the impulse vector, let $r_k(s) \in \Re^K$ be the vector response at horizon s to the k-th shock in a Cholesky-decomposition of Q_u. Then the impulse response $r_a(s)$ for a is simply given by

$$r_a(s) = \sum_{i=1}^{K} \alpha_i r_i(s).$$

For estimation consider the companion form of the state space in (2.3.4)-(2.3.5)

$$\begin{aligned} F_t =& \Phi F_{t-p} + u_t \\ X_t =& \Lambda F_t + e_t. \end{aligned}$$

To compute the impulse response vector a, let $\mathbf{a} = [a', 0_{1,K(P-1)}]'$ and compute

$$r_{a,k}(s) = (\Phi^s \mathbf{a})_k.$$

to get the impulse response of factor k to an impulse in a at horizon s. Note that $r_a(s)$ is the vector of impulse response functions of all factors to an impulse vector a at horizon s. As a second step we have to compute the impulse response functions of the single variables by combining the respective factor loading with $r_a(s)$ accordingly. This requires to compute

$$r_a^n(s) = \Lambda_n r_a(s).$$

where Λ_n is the respective n-th row vector of the factor loading matrix. We set the sign restriction on the shape of the individual impulse response functions according to the following assumption:

Assumption 1 A (contractionary) monetary policy impulse vector *is an impulse vector* $\mathbf{a}$ *so that the individual impulse response functions to* $\mathbf{a}$ *of prices and nonborrowed reserves are not positive and the impulse responses short term interest rate is positive, for a specified* *horizons s=0,...,S.*

2.4.2 Factor generalization of Cholesky Identification

For comparison we also employ the Cholesky Identification following Bernanke, Boivin and Eliasz [2005] who impose a recursive structure. Let

$$\begin{aligned} u_t &= \tilde{A} v_t \\ \tilde{A}\tilde{A}' &= Q_u \end{aligned}$$

where $\tilde{A}$ is lower triangular Cholesky factor of the factor residual VCV matrix and where the policy instrument is ordered last in the FAVAR equation. The identifying restrictions follow in parts from the factor identification which we employ as explained in Bernanke, Boivin, and Eliasz [6] which is actually independent from the identification task regarding structural shocks. The factor identification restrict the contemporaneous channel through which monetary policy shocks can effect the first $[K_c \times K_c]$ variables of our data panel which is part of the shock identification as the first variables are output related ones. Hence the contemporaneous reaction of the first $[K_c \times K_c]$ output related variables is restricted such that it does not react in the shock period. Furthermore the ordering of the panel is important. A detailed description of the identification assumptions can be found in Bernanke, Boivin, and Eliasz [6] and Stock and Watson [28].

2.5 Empirical Results

In this section we lay down the results of the paper. First we describe the data and the respective model specifications chosen. Next we present the models fit followed by the results for the convergence diagnostics. Finally we present our results for our Baseline model and 6 further Versions with the specification summarized in Table (2.1).

2.5.1 Data and Model Specification

We follow the empirical approach of Bernanke, Boivin, and Eliasz [6] who have used the data set of Stock and Watson which consists of a panel of 120 macroeconomic variables in monthly frequency transformed to induce stationarity. For comparison purposes we use a the same data span from January 1959 through August 2001. For our analysis we use an updated version as employed in Stock and Watson [28]. The data set is listed in Appendix A. The "core VAR" consists in our preferred baseline model of the policy instrument[7] and the CPI index, which is necessary to avoid misspecification as explained in the previous section. We add 4 factors, which turns out to be sufficient to capture the crucial dynamics driving the economy. We experimented with changing the number of factors, and found little information was added by increasing the dimension of the system by analyzing the marginal contribution of further factor for the explanation of the variation in the data based on R^2. For our Baseline model we report results based on a lag length of 12 where the federal funds rate and CPI are

[7] In our case the Federal Funds rate

included as factors in Y_t. We tried several versions with different lag length, but the results compared to the ones reported here do not differ much.

Table 2.1: **Different model specifications analyzed.**

Model Version	Core VAR	Identification	Restriction
Model A *[Baseline Model]*	[F_4,CPI, FFR]	SR	Monetary [M] (-): 92,96,97,98 Interest Rates [R](+): 77 Prices (-) [P]: 102,108
Model B *[Pure FAVAR]*	[FFR]	SR	see Baseline I
Model C *[Max. Restriction]*	[CPI, FFR]	SR	Baseline I + R: 78,79,80 P: 102 to 117 Spreads (-): 88,89 Output (-): 16
Model D *[Output Restr.]*	[CPI, FFR]	SR	Baseline I + Output (-): 16
Model E *[Recursive Model]*	[CPI, FFR]	Cholesky	FAVAR Ordering
Model F *[Minimal Baseline]*	[CPI, FFR]	SR	M: 98 R: 77 P: 102,108
Model G *[Recursive Pure FAVAR]*	[FFR]	Recursive	Ordering

The horizon for the sign restriction to hold is set to $H = 6$ which is based on Uhlig [31]. Not reported results confirm the tendency that the impact and intensity of the responses of restricted variables increases with the horizon H. Our different model specifications are:

1. **Model A** is the *Baseline Model* which includes as the Core VAR FFR and CPI[8].

2. **Model B** is the *Pure FAVAR*. In this model version the specification is the same as in **Model A** except that we exclude the CPI variable in the core VAR. We will see that the potential misspecification by excluding CPI is less severe when employing the sign restriction identification than in the competing recursive scheme.

3. **Model C** is the *Maximal Price Restrictions* version. It turns out that restricting prices are key to identifying monetary policy shocks. They deliver the crucial "*byte*" in the space of possible reaction of the variables and hence narrow down the space of the potential impact of monetary policy shock approximating an exact identification. In this model we restrict the maximal number of prices that serve correct responses.

4. **Model D** is the *Output* version. Some might argue that the restrictions are not sufficient because output is left unrestricted. Therefore we added here to the

[8] For the analysis of the identification of monetary policy shocks the ordering of the FAVAR equation is irrelevant in the case of sign restriction. When we identify the shocks in the recursive case the identifying assumption is that the factors are ordered first before the CPI and the policy instrument last.

Baseline Model the restriction output to react negative for some periods following the shock. Apart from that the model is identical to the first case.

5. **Model E** is the *Baseline Model with Recursive Identification*. The Specification is identical to **Baseline Model**. The identification scheme is the recursive Cholesky identification restricting the channels through which the policy instrument affects the variables in the shock period via the factors. The factors are ordered first before CPI and the policy instrument is ordered last.

6. **Model F** is the *Baseline Model with Minimal Restriction*. The Specification is identical to **Baseline Model**. The main distinction is for the restrictions imposed on the shape of the impulse response functions are according to the ones set in Uhlig [2005] for the VAR framework.

7. **Model G** is the *Baseline Model without CPI and Recursive Identification*. This version is the same as **Model E** except that we exclude CPI from the core VAR in order to analyze the impact of the misspecification discussed in section 5.

2.5.2 Model Fit

We want to assure that our methodology can represent the data in an adequate manner. In this section we want to report results that pursue to assess the models fit to the data. The first obvious check is to see how well the factors represent our panel of data series. To this end, we estimated the R^2s from regressing the respective series onto the five factors. Results are listed in the Table (2.4) below. These results give an report to what extend the individual series are driven by common components. Those with a low R^2 are more driven by idiosyncratic forces. Hence the factor model is informative in the sense that it describes the aggregate common components that we are interested in. Thus a VAR in these factors or common components should not suffer from omitted variable bias, which implies that adding individual series to the VAR in eq. (3) above does not alter the shape of any impulse response functions substantially. As already described above we also estimated the respective marginal contribution of each factor, not reported here, for the explanatory part in the variation of the data.

2.5.3 Convergence Diagnostics

In order to assure that the results are based on converged simulations, that represent the respective target distribution completely and e.g. not only some local mode we have to apply further convergence diagnostics for the simulated parameters based on the Gibbs sampler. The respective diagnostics are not a guarantee for convergence but a battery of different diagnostics that can reduce the uncertainty researchers have regarding sampled parameters. Therefore we employed some diagnostics that are widespread in the MCMC literature which are only briefly explained here and reported in Appendix C.

All our results are based on 25000 simulation draws of the Gibbs sampler of which the first 5000 were discarded as a burn-in to avoid an initial transient of the starting values that initiated the simulation. Furthermore we chose a thinning parameter of 2 so that only every second draw was kept after the burn-in in order to reduce potential autocorrelation of the sequence. Furthermore we checked the precision of the sampler by plotting the associated error bands. We furthermore checked the convergence by monitoring the sampler visually through trace plots which shows the evolvement of the sampler and helps to check whether there are jumps in the level of the respective parameter. Furthermore we plotted the first half of the kept draws against the second to check whether the sampler is still *moving* towards the target distribution or whether the whole density of the distribution is represented. Few deviation indicates convergence and vice versa [9].

2.5.4 Impulse Response Analysis

The key statistics suited to answer the question at hand are impulse response functions. Here we report the results for the different model versions as previously described and summarized in Table 1 in Appendix A. Starting with our benchmark specification **Model A** in Figure (2.1) we find a negative reaction for industrial production with a maximal impact after about a year before reverting back to pre shock levels. This is in contrast to BVAR results reporting an ambiguous effect[10]. For CPI we find a negative impact further decreasing for the entire horizon of 48 month considered. In contrast the commodity price index decreases less strong with a maximum impact after about a year before reverting back to the pre shock level a year later. Compared to CPI this is not surprising as the commodity price index reflects measures traded on the market with rather flexible prices. The federal funds rate and other short term interest rates increase on impact but revert quite soon afterwards. After about a year it swings back to a slightly lower. Other interesting variables, such as capacity utilization rate, housing starts and new orders react rather similar by decreasing for about a year before reverting. In Figure (2.5) we report results for the same specification as the benchmark case except that we also restrict industrial production to react non positively. We find the results to be quite similar. This restriction doesn't seem to change much. For comparing the differences in the results produced by the different identification schemes Figure (2.6) shows the impulse response functions of our baseline specification identified via a recursive Cholesky identification where the federal funds rate is ordered last. The most striking difference is the unreasonable high and long lasting "price puzzle". The CPI measure remains positive for about 18 month which is rather long. Even commodity prices increase on impact. Another anomaly is the increasing reaction of nonborrowed reserves which lasts for the whole horizon considered. A feature that has been analyzed in several studies because of its importance is changes in the monetary policy regime. We don't analyze that feature in greater detail as it is beyond the scope of the paper. But we provide results for the post Volcker disinflation period in order to check the fragility of the results. Figure (2.2) provides the results of both identification schemes in one plot. The crucial difference regarding the full sample results and the post Volcker disinflation period for the sign restriction approach is that the response of industrial production is ambiguous. We do not find an unambiguous decrease anymore. The federal funds rate and other short term interest rates tend to have a positive reaction for a much longer period than in the full sample case. CPI reacts less stronger

[9]Further statistics for the convergence such as the Rafetery-Lewis test, the I-statistic and more have been calculated to ensure convergence. Due to space limitations these results are available upon request.

[10]See Uhlig [31].

and reverts back after about 14 month opposed to the full sample results where CPI decrease for the whole horizon considered. The picture on the labor market is also somewhat different. Unemployment is affected stronger and for a longer period by a monetary policy shock. However employment reacts slightly after about 2 years. In comparison with results based on Cholesky identification we still find the "price puzzle". Nominal variables tend to have a stronger reaction comparing with results based on sign restriction.

2.5.5 Forecast Error Variance Decomposition

An important question after having identified monetary policy shocks is how much of the variation it explains. Results are provided by Figure the appendix for the sign restriction case. For comparison purposes we also provide results for the case of Cholesky identification. According to the median estimates we find that a random shift in monetary policy explains about 12%-15% in the variation of industrial production and less than about 20% for most variables except for prices and short term interest rates. These results are amongst others consistent with Sims [24],Sims and Zha [27], Uhlig [31] and Amir Ahmadi and Ritschl [1].

2.6 Discussion

The impulse responses presented serve the indication that, for the researcher interested in measuring the effects of a shock to monetary policy, identifying restrictions that are consistent with the conventional wisdom, such as the proposed sign restrictions are decisive for the results to look reasonable. We have shown that our identification strategy is particularly well suited in a factor based model such as the employed BFAVAR. When comparing the results from the Gibbs sampling approach identified with the recursive identification scheme it is quite evident that the results seem more reasonable w.r.t. to unreasonable anomalies present such as the "price puzzle".

2.6.1 Advantage Towards BVAR.

For Compactness we will briefly summarize the key arguments why factor based models are favorable compared to (B)VARs. The parsimoniously exploitation of a large set of disaggregated macroeconomic data through the dimension reduction methodology of BFAVARs avoids the omitted variable bias and overcomes the degree-of-freedom problem which are inherent drawbacks of the (B)VAR approach. Hence we do not have to restrict ourselves to a small scale VAR that apparently serves an unrealistic and fragmentary picture of the economy and the central bankers information set prior to taking decision. Hence small scale VARs might be a pitfall of biased inference. Furthermore the luxury position of large panels of macroeconomic data allows to check model validation and the robustness of the respective identification approaches applied due to the fact that a crucial amount of important data is considered. Hence we can have a much broader picture of the economy's reaction after structural shocks. On account of this we can have a thorough picture of the impact of shocks by considering and analyzing more variables of interest, other than the standard VAR indicators. We furthermore do not restrict ourselves to rely on single indicators that can impossibly reflect a whole economic concept. In such cases, inference might strongly depend the choice of such indicators. Through the methodology we let the data decide which weight to give to the single disaggregated series without excluding them a priori. Another important advantage is that the crucial dynamics of the data rich economy are captured more appropriate. Furthermore we find as an empirical result that the BFAVAR approach provides significant real effects whereas in the (B)VAR approach one comes up with insignificant results where no statement is possible. The apparent reason might be the lack of information included in the (B)VAR framework.

2.6.2 Advantage towards Nonparametric FAVAR.

Admittedly one must say that the Bayesian approach by applying the Gibbs sampler in combination with the sign restriction is a cumbersome and time consuming procedure which is a disadvantage towards the nonparametric two step estimation procedure applied by BBE. The Two-Step estimation procedure delivers good results and is efficient regarding the computation procedure but cannot exploit the fully parametric factor structure with a lag polynomial which we evaluate of great relevance. Dynamic Principal component analysis and other estimation procedures for classical DFM's are convenient ways to reduce dimensions due to their computational efficiency and easy implementation. Nevertheless they lack on the possibility to exploit the factor structure. For that reason we prefer to take the Bayesian approach by estimating the model in the state space representation in order to impose an explicit parametrical model. The State space representation moreover serves the possibility to let the dynamics of the state equation as unrestricted as possible and furthermore introduce sophisticated priors that can account for crucial interaction between the driving factors and forces of the economy. Hence we can be even more precise by capturing more realistically important comovements.

2.6.3 Sign Restriction vs. Recursive Identification in a BFAVAR approach.

By comparing the results of the recursive identification scheme and the sign restriction approach it turns out that the results of the latter are associated with less uncertainty as can be seen from the error bands. This is an enticing promise insofar that an economically desirable identification scheme that is "weak" with respect to the restrictions and structure imposed delivers reasonable results.

2.7 Conclusion

In this paper we propose to estimate the effects of monetary policy shocks in a BFAVAR framework identified via sign restriction. We estimate the model from a Bayesian perspective relying on MCMC methods. To infer the effects of a contractionary shock to monetary policy we impose sign restriction on the impulses responses on many prices, monetary aggregates, spreads and short term interest rates. More restrictions can be set in factor based models which is promising as the robustness of identification can be monitored and identifying assumption, tight to economic theory can be imposed broadly. Our identification strategy outperforms the recursive identification in the BFAVAR framework hence we conclude that the combination of the BFAVAR model and a reasonable identification scheme are crucial for a successful identification. Furthermore, sign restrictions hold independent from subsamples considered unlike the Cholesky case where results and inference are sensitive sometimes contradicting economic theory. We find that after a contractionary monetary policy shock a negative response of output but modest in size. This effect vanishes when we consider the post Volcker disinflation period. Monetary policy shocks account for a small fraction of the variance for the forecast revision of output.

Appendix A

Appendix A.1: Tables

Table (2.2) is taken from Stock and Watson [28] and lists the short name of each series, its mnemonic (the series label used in the source database), the transformation applied to the series, and a brief data description. All series are from the Global Insights Basic Economics Database, unless the source is listed (in parentheses) as TCB (The Conference Board's Indicators Database) or AC (author's calculation based on Global Insights or TCB data). In the transformation column, 1 denotes the level of the series, 2 denotes the first difference of the series, 4 denotes logarithm, 5 and 6 denote the first and second difference of the logarithm respectively.

Pos	Shortname	Mnemonics	TC	Description
1	PI	a0m052	5	Personal income (AR, bil. chain 2000 $)
2	PI less transfers	A0M051	5	Personal income less transfer payments (AR, bil. chain 2000 $)
3	Consumption	A0M224R	5	Real Consumption (AC) A0m224/gmdc
4	M&T sales	A0M057	5	Manufacturing and trade sales (mil. Chain 1996 $)
5	Retail sales	A0M059	5	Sales of retail stores (mil. Chain 2000 $)
6	IP: total	IPS10	5	INDUSTRIAL PRODUCTION INDEX - TOTAL INDEX
7	IP: products	IPS11	5	INDUSTRIAL PRODUCTION INDEX - PRODUCTS, TOTAL
8	IP: final prod	IPS299	5	INDUSTRIAL PRODUCTION INDEX - FINAL PRODUCTS
9	IP: cons gds	IPS12	5	INDUSTRIAL PRODUCTION INDEX - CONSUMER GOODS
10	IP: cons dble	IPS13	5	INDUSTRIAL PRODUCTION INDEX - DURABLE CONSUMER GOODS
11	iIP:cons nondble	IPS18	5	INDUSTRIAL PRODUCTION INDEX - NONDURABLE CONSUMER GOODS
12	IP:bus eqpt	IPS25	5	INDUSTRIAL PRODUCTION INDEX - BUSINESS EQUIPMENT
13	IP: matls	IPS32	5	INDUSTRIAL PRODUCTION INDEX - MATERIALS
14	IP: dble mats	IPS34	5	INDUSTRIAL PRODUCTION INDEX - DURABLE GOODS MATERIALS
15	IP:nondble mats	IPS38	5	INDUSTRIAL PRODUCTION INDEX - NONDURABLE GOODS MATERIALS
16	IP: mfg	IPS43	5	INDUSTRIAL PRODUCTION INDEX - MANUFACTURING (SIC)
17	IP: res util	IPS307	5	INDUSTRIAL PRODUCTION INDEX - RESIDENTIAL UTILITIES
18	IP: fuels	IPS306	5	INDUSTRIAL PRODUCTION INDEX - FUELS
19	NAPM prodn	PMP	1	NAPM PRODUCTION INDEX (PERCENT)
20	Cap util	A0m082	2	Capacity Utilization (Mfg)
21	Help wanted indx	LHEL	2	INDEX OF HELP-WANTED ADVERTISING IN NEWSPAPERS (1967=100;SA)
22	Help wanted/emp	LHELX	2	EMPLOYMENT: RATIO; HELP-WANTED ADS:NO. UNEMPLOYED CLF
23	Emp CPS total	LHEM	5	CIVILIAN LABOR FORCE: EMPLOYED, TOTAL (THOUS.,SA)
24	Emp CPS nonag	LHNAG	5	CIVILIAN LABOR FORCE: EMPLOYED, NONAGRIC.INDUSTRIES (THOUS.,SA)
25	U: all	LHUR	2	UNEMPLOYMENT RATE: ALL WORKERS, 16 YEARS & OVER (%,SA)
26	U: mean duration	LHU680	2	UNEMPLOY.BY DURATION: AVERAGE(MEAN)DURATION IN WEEKS (SA)
27	U < 5 wks	LHU5	5	UNEMPLOY.BY DURATION: PERSONS UNEMPL.LESS THAN 5 WKS (THOUS.,SA)
28	U 5-14 wks	LHU14	5	UNEMPLOY.BY DURATION: PERSONS UNEMPL.5 TO 14 WKS (THOUS.,SA)
29	U 15+ wks	LHU15	5	UNEMPLOY.BY DURATION: PERSONS UNEMPL.15 WKS + (THOUS.,SA)
30	U 15-26 wks	LHU26	5	UNEMPLOY.BY DURATION: PERSONS UNEMPL.15 TO 26 WKS (THOUS.,SA)
31	U 27+ wks	LHU27	5	UNEMPLOY.BY DURATION: PERSONS UNEMPL.27 WKS + (THOUS,SA)
32	UI claims	A0M005	5	Average weekly initial claims, unemploy. insurance (thous.)
33	Emp: total	CES002	5	EMPLOYEES ON NONFARM PAYROLLS - TOTAL PRIVATE
34	Emp: gds prod	CES003	5	EMPLOYEES ON NONFARM PAYROLLS - GOODS-PRODUCING
35	Emp: mining	CES006	5	EMPLOYEES ON NONFARM PAYROLLS - MINING
36	Emp: const	CES011	5	EMPLOYEES ON NONFARM PAYROLLS - CONSTRUCTION
37	Emp: mfg	CES015	5	EMPLOYEES ON NONFARM PAYROLLS - MANUFACTURING
38	Emp: dble gds	CES017	5	EMPLOYEES ON NONFARM PAYROLLS - DURABLE GOODS
39	Emp: nondbles	CES033	5	EMPLOYEES ON NONFARM PAYROLLS - NONDURABLE GOODS
40	Emp: services	CES046	5	EMPLOYEES ON NONFARM PAYROLLS - SERVICE-PROVIDING
41	Emp: TTU	CES048	5	EMPLOYEES ON NONFARM PAYROLLS - TRADE, TRANSPORTATION
42	Emp: wholesale	CES049	5	EMPLOYEES ON NONFARM PAYROLLS - WHOLESALE TRADE
43	Emp: retail	CES053	5	EMPLOYEES ON NONFARM PAYROLLS - RETAIL TRADE
44	Emp: FIRE	CES088	5	EMPLOYEES ON NONFARM PAYROLLS - FINANCIAL ACTIVITIES
45	Emp: Govt	CES140	5	EMPLOYEES ON NONFARM PAYROLLS - GOVERNMENT
46	Emp-hrs nonag	A0M048	5	Employee hours in nonag. establishments (AR, bil. hours)
47	Avg hrs	CES151	1	AVERAGE WEEKLY HOURS OF PRODUCTION OR NONSUPERVISORY WKRS
48	Overtime: mfg	CES155	2	AVERAGE WEEKLY HOURS OF PRODUCTION OR NONSUPERVISORY
49	Avg hrs: mfg	aom001	1	Average weekly hours, mfg. (hours)
50	NAPM empl	PMEMP	1	NAPM EMPLOYMENT INDEX (PERCENT)
51	HStarts: Total	HSFR	4	HOUSING STARTS:NONFARM(1947-58);TOTAL FARM&NONFARM(1959-)
52	HStarts: NE	HSNE	4	HOUSING STARTS:NORTHEAST (THOUS.U.)S.A.
53	HStarts: MW	HSMW	4	HOUSING STARTS:MIDWEST(THOUS.U.)S.A.
54	HStarts: South	HSSOU	4	HOUSING STARTS:SOUTH (THOUS.U.)S.A.
55	HStarts: West	HSWST	4	HOUSING STARTS:WEST (THOUS.U.)S.A.
56	BP: total	HSBR	4	HOUSING AUTHORIZED: TOTAL NEW PRIV HOUSING UNITS (THOUS.,SAAR)

Pos	Shortname	Mnemonics	TC	Description
57	BP: NE	HSBNE	4	HOUSES AUTHORIZED BY BUILD. PERMITS:NORTHEAST(THOU.U.)S.A
58	BP: MW	HSBMW	4	HOUSES AUTHORIZED BY BUILD. PERMITS:MIDWEST(THOU.U.)S.A.
59	BP: South	HSBSOU	4	HOUSES AUTHORIZED BY BUILD. PERMITS:SOUTH(THOU.U.)S.A.
60	BP: West	HSBWST	4	HOUSES AUTHORIZED BY BUILD. PERMITS:WEST(THOU.U.)S.A.
61	PMI	PMI	1	PURCHASING MANAGERS' INDEX (SA)
62	NAPM new ordrs	PMNO	1	NAPM NEW ORDERS INDEX (PERCENT)
63	NAPM vendor del	PMDEL	1	NAPM VENDOR DELIVERIES INDEX (PERCENT)
64	NAPM Invent	PMNV	1	NAPM INVENTORIES INDEX (PERCENT)
65	Orders: cons gds	A0M008	5	Mfrs' new orders, consumer goods and materials (bil. chain 1982 $)
66	Orders: dble gds	A0M007	5	Mfrs' new orders, durable goods industries (bil. chain 2000 $)
67	Orders: cap gds	A0M027	5	Mfrs' new orders, nondefense capital goods (mil. chain 1982 $)
68	Unf orders: dble	A1M092	5	Mfrs' unfilled orders, durable goods indus. (bil. chain 2000 $)
69	M&T invent	A0M070	5	Manufacturing and trade inventories (bil. chain 2000 $)
70	M&T invent/sales	A0M077	2	Ratio, mfg. and trade inventories to sales (based on chain 2000 $)
71	M1	FM1	5	MONEY STOCK: M1(CURR,TRAV.CKS,DEM DEP,OTHER CK'ABLE DEP)(BIL$,SA)
72	M2	FM2	5	MONEY STOCK:M2(M1+O'NITE RPS,EURO$,G/P&B/D MMMFS&SAV&SM
73	M3	FM3	5	MONEY STOCK: M3(M2+LG TIME DEP,TERM RP'S&INST ONLY MMMFS)(BIL$,SA)
74	M2 (real)	FM2DQ	5	MONEY SUPPLY - M2 IN 1996 DOLLARS (BCI)
75	MB	FMFBA	5	MONETARY BASE, ADJ FOR RESERVE REQUIREMENT CHANGES(MIL$,SA)
76	Reserves tot	FMRRA	5	DEPOSITORY INST RESERVES:TOTAL,ADJ FOR RESERVE REQ CHGS(MIL$,SA)
77	Reserves nonbor	FMRNBA	5	DEPOSITORY INST RESERVES:NONBORROWED,ADJ RES REQ CHGS(MIL$,SA)
78	C&I loans	FCLNQ	5	COMMERCIAL & INDUSTRIAL LOANS OUSTANDING IN 1996 DOLLARS (BCI)
79	C&I loans	FCLBMC	1	WKLY RP LG COM'L BANKS:NET CHANGE COM'L & INDUS LOANS(BIL$,SAAR)
80	Cons credit	CCINRV	5	CONSUMER CREDIT OUTSTANDING - NONREVOLVING(G19)
81	Inst cred/PI	A0M095	2	Ratio, consumer installment credit to personal income (pct.)
82	S&P 500	FSPCOM	5	S&P'S COMMON STOCK PRICE INDEX: COMPOSITE (1941-43=10)
83	S&P: indust	FSPIN	5	S&P'S COMMON STOCK PRICE INDEX: INDUSTRIALS (1941-43=10)
84	S&P div yield	FSDXP	2	S&P'S COMPOSITE COMMON STOCK: DIVIDEND YIELD (% PER ANNUM)
85	S&P PE ratio	FSPXE	5	S&P'S COMPOSITE COMMON STOCK: PRICE-EARNINGS RATIO (%,NSA)
86	FedFunds	FYFF	2	INTEREST RATE: FEDERAL FUNDS (EFFECTIVE) (% PER ANNUM,NSA)
87	Commpaper	CP90	2	Cmmercial Paper Rate (AC)
88	3 mo T-bill	FYGM3	2	INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,3-MO.(% PER ANN,NSA)
89	6 mo T-bill	FYGM6	2	INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,6-MO.(% PER ANN,NSA)
90	1 yr T-bond	FYGT1	2	INTEREST RATE: U.S.TREASURY CONST MATURITIES,1-YR.(% PER ANN,NSA)
91	5 yr T-bond	FYGT5	2	INTEREST RATE: U.S.TREASURY CONST MATURITIES,5-YR.(% PER ANN,NSA)
92	10 yr T-bond	FYGT10	2	INTEREST RATE: U.S.TREASURY CONST MATURITIES,10-YR.(% PER ANN,NSA)
93	Aaabond	FYAAAC	2	BOND YIELD: MOODY'S AAA CORPORATE (% PER ANNUM)
94	Baa bond	FYBAAC	2	BOND YIELD: MOODY'S BAA CORPORATE (% PER ANNUM)
95	CP-FF spread	scp90	1	cp90-fyff
96	3 mo-FF spread	sfygm3	1	fygm3-fyff
97	6 mo-FF spread	sFYGM6	1	fygm6-fyff
98	1 yr-FF spread	sFYGT1	1	fygt1-fyff
99	5 yr-FFspread	sFYGT5	1	fygt5-fyff
100	10yr-FF spread	sFYGT10	1	fygt10-fyff
101	Aaa-FF spread	sFYAAAC	1	fyaaac-fyff
102	Baa-FF spread	sFYBAAC	1	fybaac-fyff
103	Ex rate: avg	EXRUS	5	UNITED STATES;EFFECTIVE EXCHANGE RATE(MERM)(INDEX NO.)
104	Ex rate: Switz	EXRSW	5	FOREIGN EXCHANGE RATE: SWITZERLAND (SWISS FRANC PER U.S.$)
105	Ex rate: Japan	EXRJAN	5	FOREIGN EXCHANGE RATE: JAPAN (YEN PER U.S.$)
106	Ex rate: UK	EXRUK	5	FOREIGN EXCHANGE RATE: UNITED KINGDOM (CENTS PER POUND)
107	EX rate: Canada	EXRCAN	5	FOREIGN EXCHANGE RATE: CANADA (CANADIAN $ PER U.S.$)
108	PPI: fin gds	PWFSA	5	PRODUCER PRICE INDEX: FINISHED GOODS (82=100,SA)
109	PPI: cons gds	PWFCSA	5	PRODUCER PRICE INDEX:FINISHED CONSUMER GOODS (82=100,SA)
110	PPI: int matls	PWIMSA	5	PRODUCER PRICE INDEX:INTERMED MAT.SUPPLIES & COMPONENTS
111	PPI: crude matls	PWCMSA	5	PRODUCER PRICE INDEX:CRUDE MATERIALS (82=100,SA)
112	Commod: spot price	PSCCOM	5	SPOT MARKET PRICE INDEX:BLS & CRB: ALL COMMODITIES(1967=100)
113	Sens matls price	PSM99Q	5	INDEX OF SENSITIVE MATERIALS PRICES (1990=100)(BCI-99A)
114	NAPM com price	PMCP	1	NAPM COMMODITY PRICES INDEX (PERCENT)
115	CPI-U: all	PUNEW	5	CPI-U: ALL ITEMS (82-84=100,SA)
116	CPI-U: apparel	PU83	5	CPI-U: APPAREL & UPKEEP (82-84=100,SA)
117	CPI-U: transp	PU84	5	CPI-U: TRANSPORTATION (82-84=100,SA)
118	CPI-U: medical	PU85	5	CPI-U: MEDICAL CARE (82-84=100,SA)
119	CPI-U: comm.	PUC	5	CPI-U: COMMODITIES (82-84=100,SA)
120	CPI-U: dbles	PUCD	5	CPI-U: DURABLES (82-84=100,SA)
121	CPI-U: services	PUS	5	CPI-U: SERVICES (82-84=100,SA)
122	CPI-U: ex food	PUXF	5	CPI-U: ALL ITEMS LESS FOOD (82-84=100,SA)
123	CPI-U: ex shelter	PUXHS	5	CPI-U: ALL ITEMS LESS SHELTER (82-84=100,SA)
124	CPI-U: ex med	PUXM	5	CPI-U: ALL ITEMS LESS MEDICAL CARE (82-84=100,SA)
125	PCE defl	GMDC	5	PCE,IMPL PR DEFL:PCE (1987=100)
126	PCE defl: dlbes	GMDCD	5	PCE,IMPL PR DEFL:PCE; DURABLES (1987=100)
127	PCE defl: nondble	GMDCN	5	PCE,IMPL PR DEFL:PCE; NONDURABLES (1996=100)
128	PCE defl: services	GMDCS	5	PCE,IMPL PR DEFL:PCE; SERVICES (1987=100)
129	AHE: goods	CES275	5	AVERAGE HOURLY EAR. OF PRODUCTION OR NONSUPERVISORY
130	AHE: const	CES277	5	AVERAGE HOURLY EAR. OF PRODUCTION OR NONSUPERVISORY WRKRS
131	AHE: mfg	CES278	5	AVERAGE HOURLY EARN. OF PRODUCTION OR NONSUPERVISORY PRIVATE
132	Consumer expect	HHSNTN	2	U. OF MICH. INDEX OF CONSUMER EXPECTATIONS(BCD-83)

Table 2.3: **Different model specifications analyzed.**

Model Version	Core VAR	Identification	Restriction
Model A *[Baseline Model]*	[CPI, FFR]	SR	Monetary [M] (-): 92,96,97,98 Interest Rates [R](+): 77 Prices (-) [P]: 102,108
Model B *[Pure FAVAR]*	[FFR]	SR	see Baseline I
Model C *[Max. Restriction]*	[CPI, FFR]	SR	Baseline I + R: 78,79,80 P: 102 to 117 Spreads (-): 88,89 Output (-): 16
Model D *[Output Restr.]*	[CPI, FFR]	SR	Baseline I + Output (-): 16
Model E *[Recursive Model]*	[CPI, FFR]	Recursive	FAVAR Ordering
Model F *[Minimal Baseline]*	[CPI, FFR]	SR	M: 98 R: 77 P: 102,108
Model G *[Recursive Pure FAVAR]*	[FFR]	Recursive	Ordering

Table 2.4: **Share of Variance Explained by Common Factors.**

Variable	R^2	**Variable**	R^2	**Variable**	R^2
IPE	0,31	LPHRM	0,38	SFYGM	0,83
IPI	0,19	LPMOSA	0,36	SFYGT	0,92
IPM	0,23	Employment	0,91	SFYGT	0,93
IPMD	0,22	Pers Cons	0,07	SFYGT	1,00
IPMND	0,10	Dur Cons	0,03	SFYAAAC	0,99
IPMFG	0,30	Non Dor Cons	0,04	SFYBAAC	1,00
IPD	0,27	GMCSQ	0,04	M	0,99
IPN	0,17	GMCANQ	0,01	M	0,13
IPMIN	0,03	Housing Starts	0,56	M	0,05
IPUT	0,03	HSNE	0,45	FM DQ	0,07
IP	0,70	HSMW	0,62	M Base	0,31
CapUtilRate	0,71	HSSOU	0,52	FMRRA	0,17
PMI	0,96	HSWST	0,36	NBR	0,04
PMP	0,88	HSBR	0,58	FCLNQ	0,07
GMPYQ	0,09	HMOB	0,51	FCLBMC	0,17
GMYXPQ	0,27	PMNV	0,67	CCINRV	0,07
LHEL	0,28	New Order	0,87	CommPI	0,21
LHELX	0,81	PMDEL	0,67	PWSFA	0,65
LHEM	0,13	MOCMQ	0,08	PWFCSA	0,45
LHNAG	0,16	MSONDQ	0,02	PWIMSA	0,41
Uneployment	0,81	FSNCOM	0,06	PWCMSA	0,37
LHU	0,89	FSPCOM	0,06	PSMQ	0,11
LHU	0,90	FSPIN	0,05	CPI	1,00
LHU	0,95	FSPCAP	0,04	PU	0,31
LHU	0,98	FSPUT	0,06	PU	0,19
LHU	0,97	Dividends	0,66	PU	0,46
LPNAG	0,46	FSPXE	0,58	PUC	0,39
LP	0,46	EXRSW	0,03	PUCD	0,78
LPGD	0,47	XR YEN	0,06	PUS	0,41
LPMI	0,04	EXRUK	0,03	PUXF	0,58
LPCC	0,11	EXRCAN	0,03	PUXHS	0,77
LPEM	0,52	FFR	1,00	PUXM	0,80
LPED	0,45	m TreasBills	0,98	Earnings (avg/h)	0,89
LPEN	0,37	FYGM	0,98	LEHM	0,06
LPSP	0,27	FYGT	0,99	Cons Expec	0,12
LPTU	0,02	y TreasBond	0,99	Real FFR	0,60
LPT	0,29	FYGT	1,00		
LPFR	0,13	FYAAAC	1,00		
LPS	0,17	FYBAAC	1,00		
LPGOV	0,13	SFYGM	1,00		

Appendix A.2: Figures

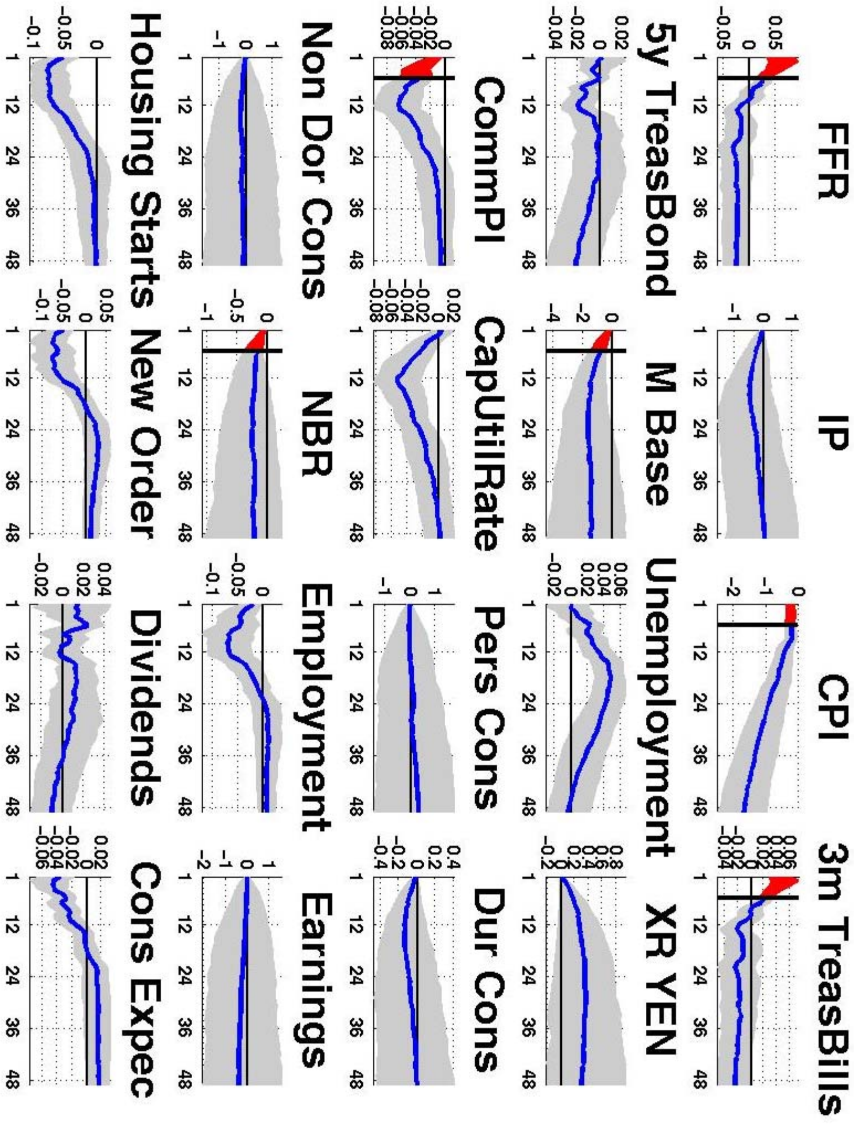


Figure 2.1: **Impulse Responses for** ***MODEL A: Baseline model***

Here we report the impulse response functions of the **I. Baseline Model**. Those series in the plot that were restricted are visualized through the red shaded area. The horizon up to which the sign restrictions were imposed to hold are denoted by the horizontal black line. We observe some real effects most notably for industrial production. However the effect is short-lived and soon after few months the impulse response is associated with increasing uncertainty along the horizon.

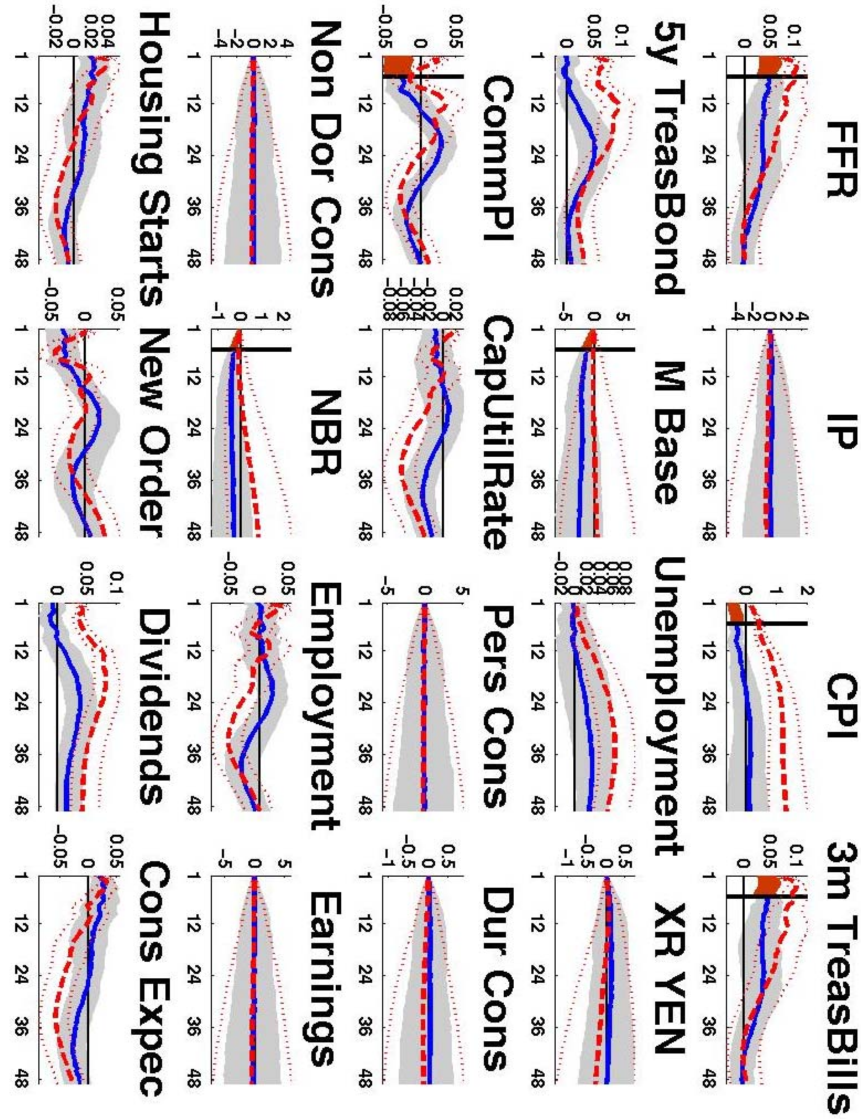


Figure 2.2: **Impulse Responses for *Post-Volcker period***
Here we compare the results for the post volcker period identified with sign restriction (blue solid line with 68 % probability bands via the grey shaded area) and Cholesky identification (red solid line represents the median and the red dotted represent the 16% and 84% percentiles respectively). Clearly for the Cholesky identification the "prize puzzle" is present and even much stronger compared the full sample period. For the CPI indicator it is even increasing for a long horizon and shows a sign of persistence. This Cholesky identification implies stronger reaction of the interest rates at different maturities and a stronger reaction of the labor market.

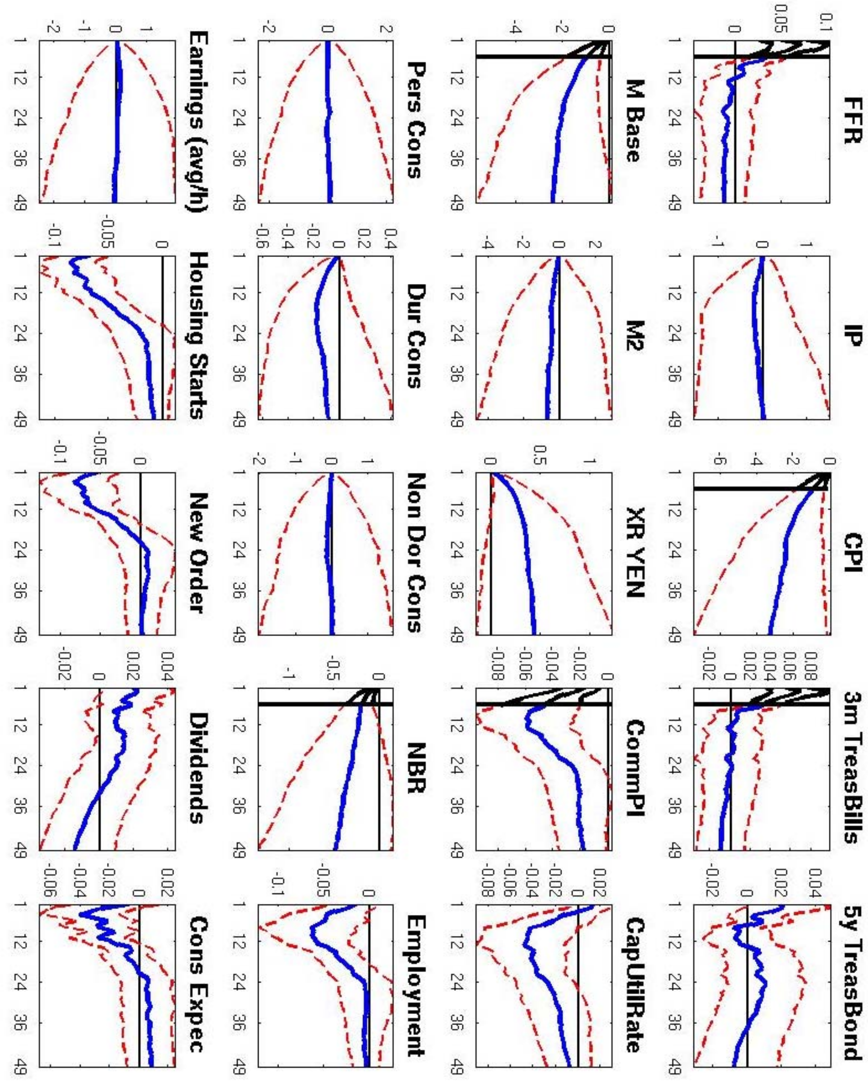


Figure 2.3: **Impulse response functions of II. Pure FAVAR**

This amounts to only consider FFR as a Factor and excluding CPI out of the core VAR. Those series in the plot that were restricted are visualized in black and a horizontal line is added to the horizon where the restriction was active.

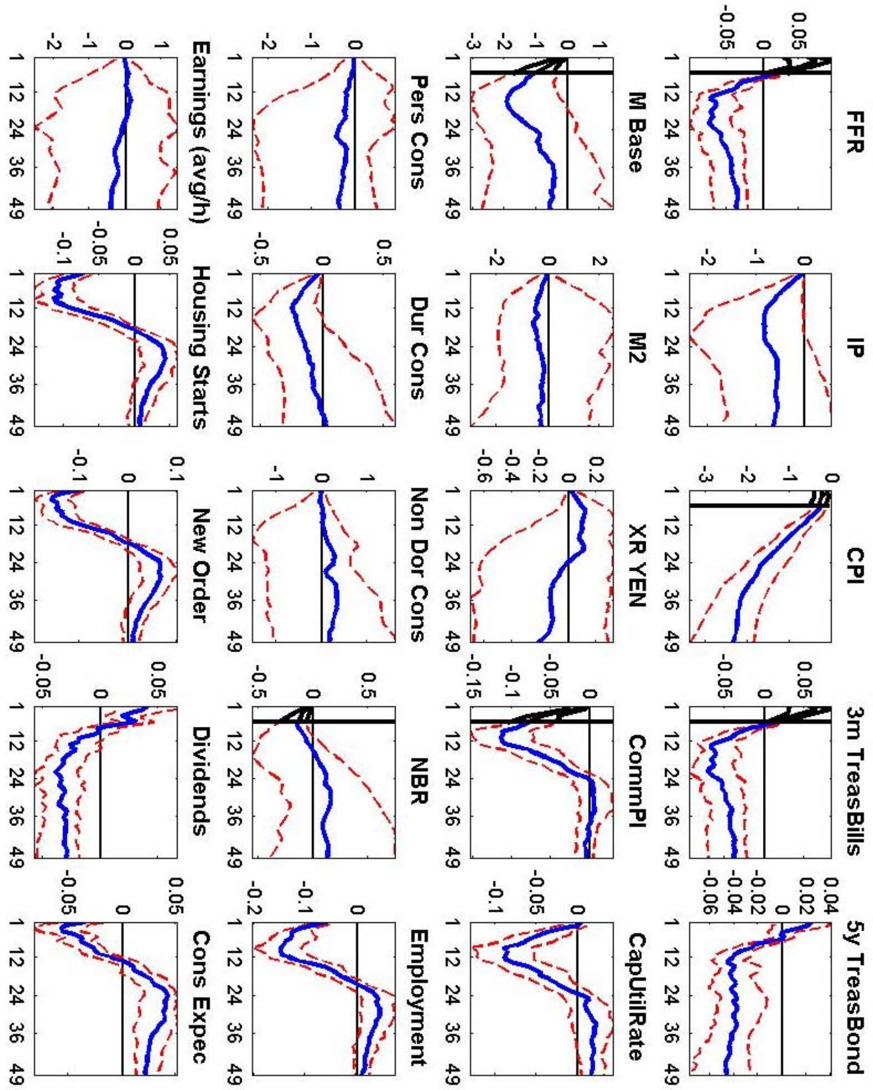


Figure 2.4: **Impulse Responses for *MODEL C: Maximal sign restrictions***

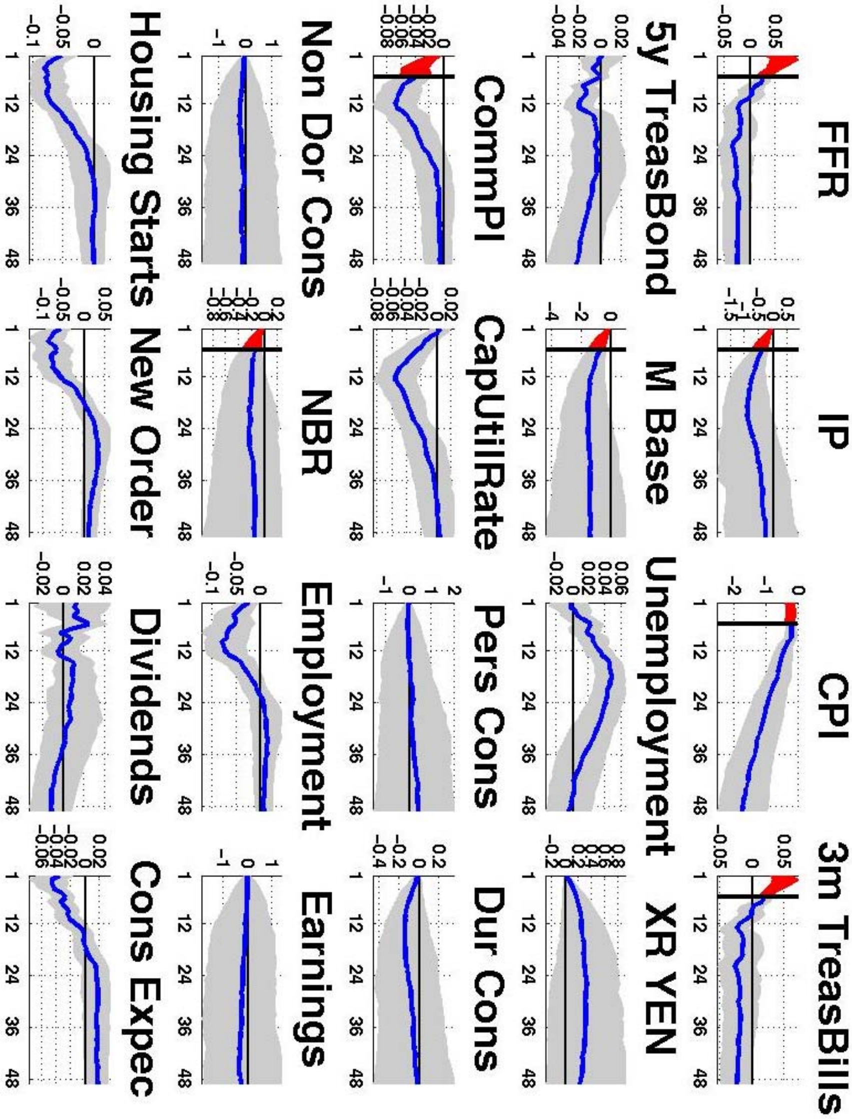


Figure 2.5: **Impulse Responses for *MODEL D: Output Model***

In this model version we additionally impose the restriction on the Industrial production that was left agnostically unrestricted to analyze its impact. The difference to the baseline model is that the reaction of output itself is slightly stronger negative with less uncertainty for a longer period of time. The remaining indicators are unaffected.

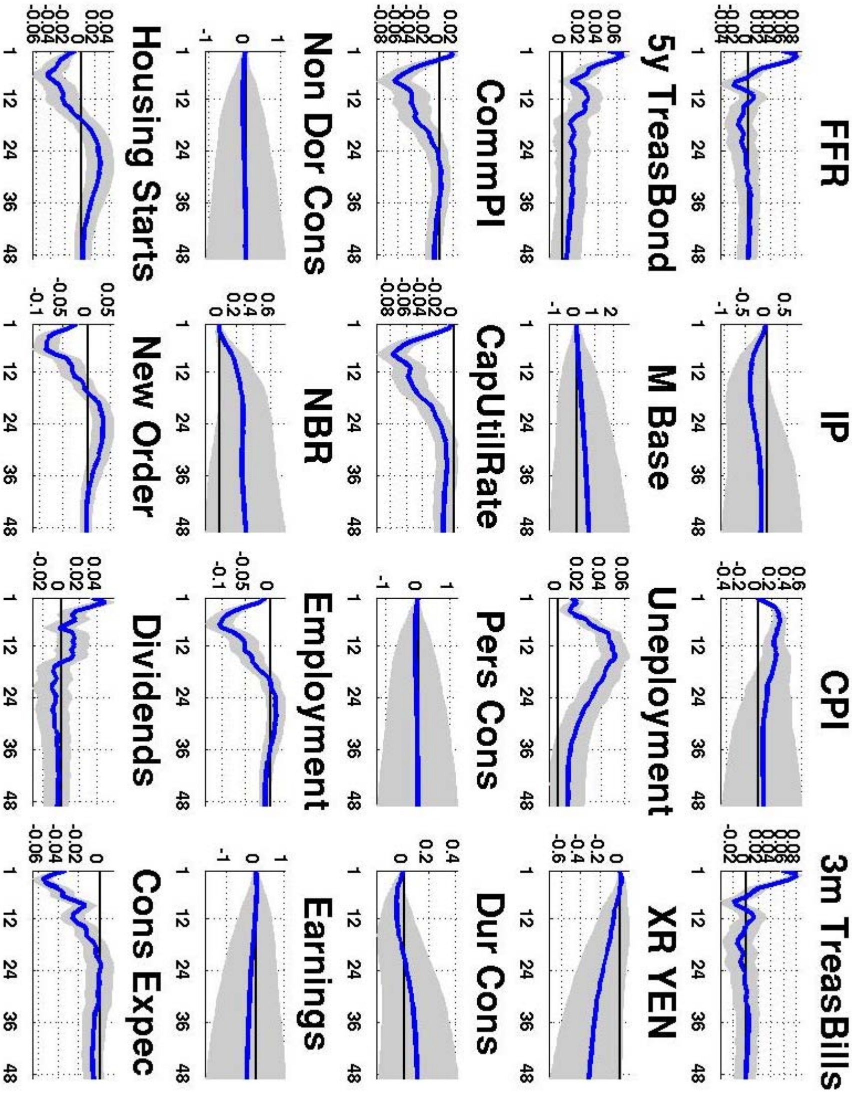


Figure 2.6: **Impulse Responses for *Model E: Baseline Cholesky***

Here the set up is exactly the same as in the baseline case of Model A however the identification is employed with the recursive Cholesky identification. The key differences in the results are the clear price puzzles in CPI and commodity price index. The latter shows how flawed this identification schemes is. The commodity price index was originally included to overcome the price puzzle which is obviously not the case here. For CPI the price puzzle holds up to two years.

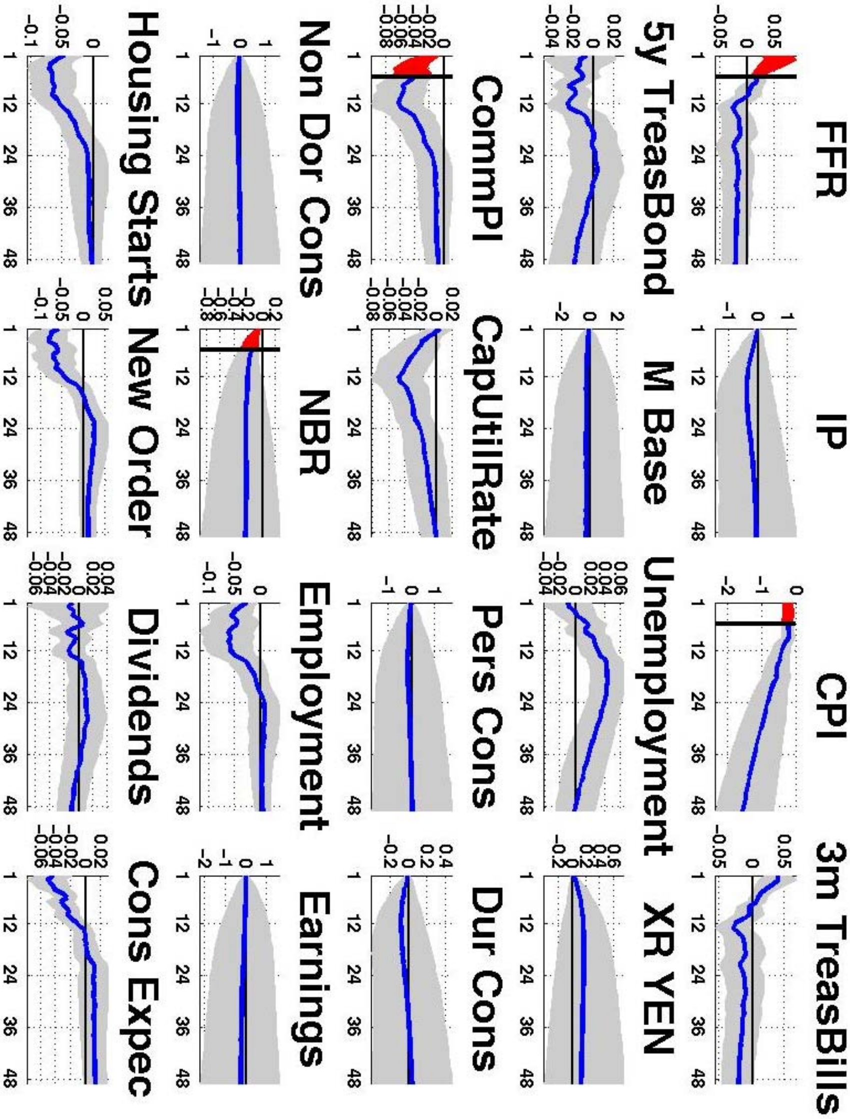


Figure 2.7: **Impulse Responses for *Model F: Baseline Minimal SR***

This figure shows the same results as in the baseline case except that we impose the minimal restrictions as proposed in Uhlig [2005]. Results are similar to the baseline case except for a slightly higher degree of uncertainty around the reaction of industrial production.

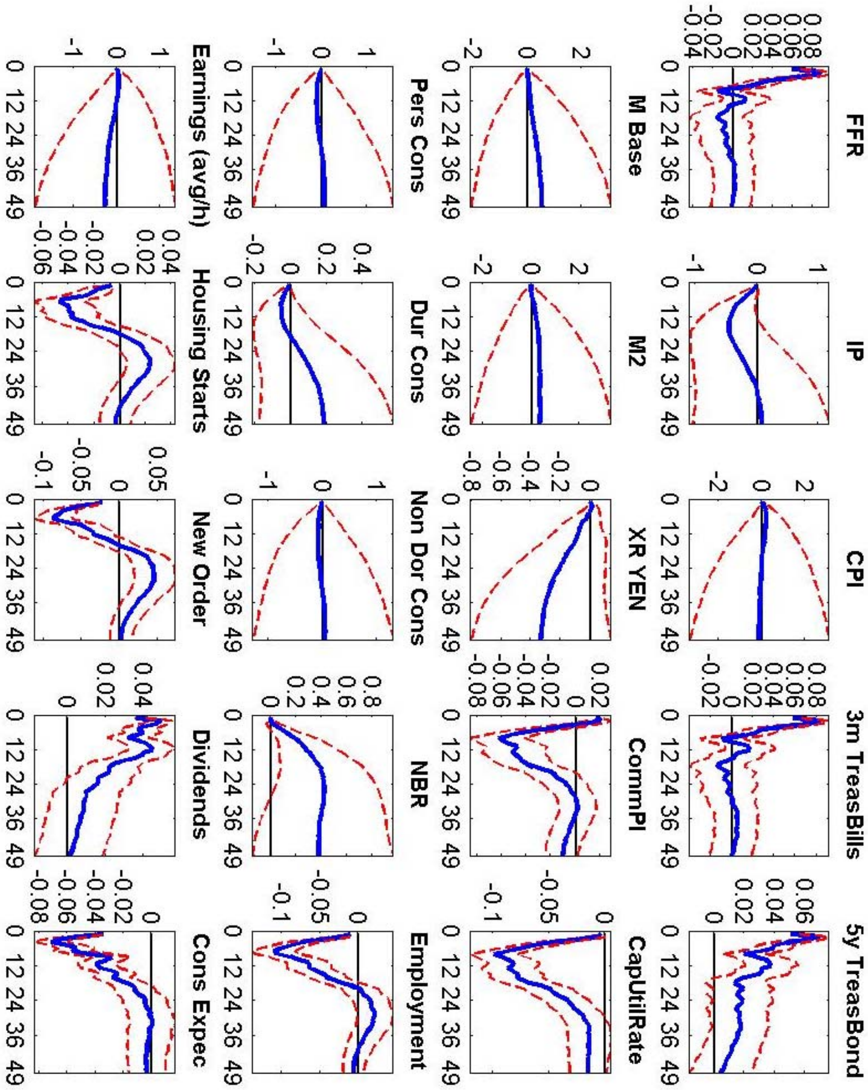


Figure 2.8: **Impulse Responses for *Model G: Baseline w/o CPI Cholesky***

This plot reports results based on the Cholesky identification where in the core VAR the CPI indicator is not included. Interestingly the strong price puzzle vanishes to the extend that a positive response is as likely as a negative. This might be due to the fact that it is less well explained by the factors. However the price puzzle in the commodity price index is still present.

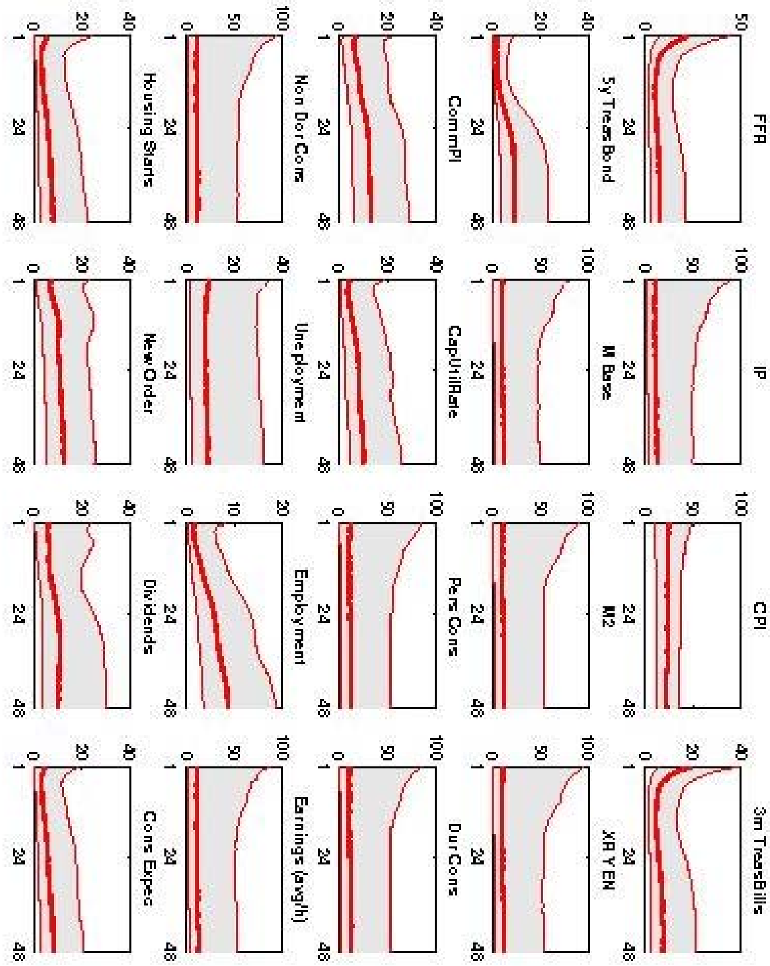


Figure 2.9: **Forecast Error Variance Decomposition** ***Model A: Baseline Model*** **identified via sign restriction.**

This plot shows the forecast error variance decomposition of a contractionary monetary policy shock in the baseline case of model A. As expected a contractionary monetary policy mostly contributes to the variation in interest rates at the short horizon of around 25 % quickly going down after 3-4 month to around 10 %. This is much lower than in the case of Cholesky identification as can be seen in the next plot. The contractionary monetary policy shocks account on average only for a low fraction in the variation of industrial production and other output indicators of around 10%-12% over the entire horizon.

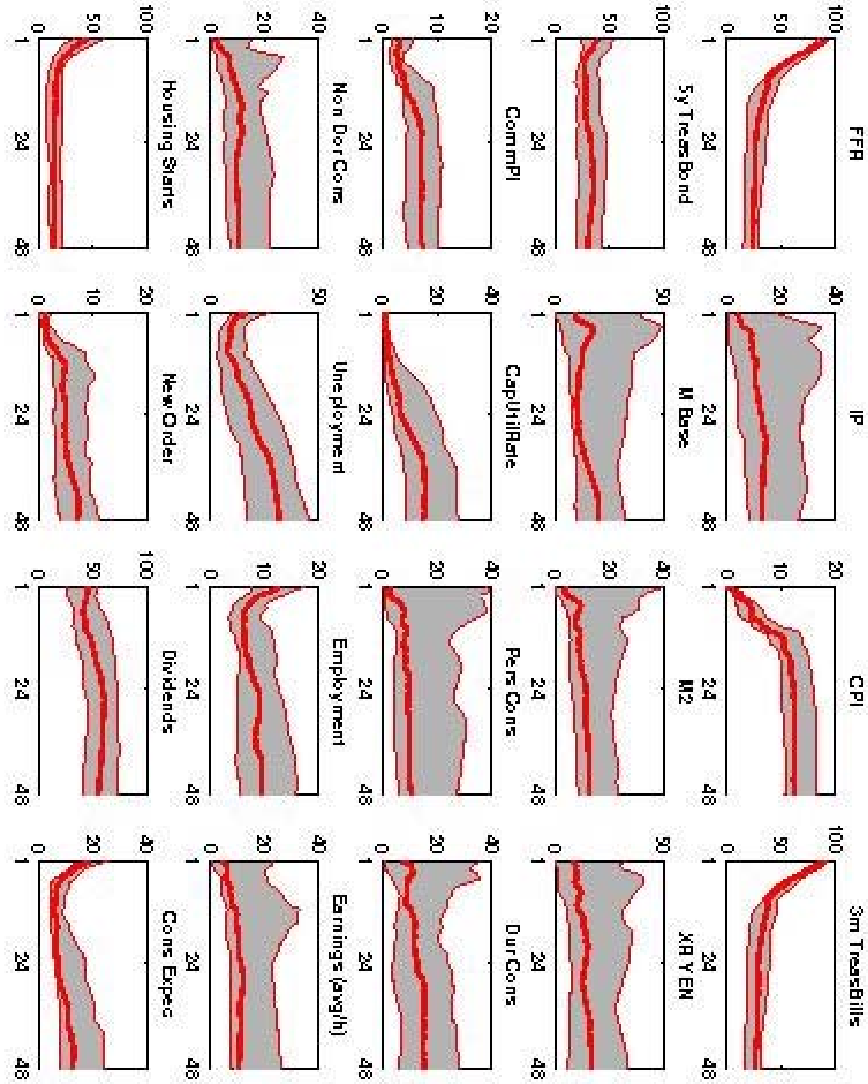


Figure 2.10: **Forecast Error Variance Decomposition** ***Model A: Baseline Model*** **identified via Cholesky.**
This plot shows the forecast error variance decomposition of baseline model with Cholesky identification.

3 Monetary Policy During the Great Depression

Joint with Albrecht Ritschl

The prominent role of monetary policy in the U.S. interwar depression has been conventional wisdom since Friedman and Schwartz [15]. Doubts surrounding interwar GDP estimates would call into question conventional monetary VAR evidence. This paper therefore attempts to capture the pertinent dynamics through a factor-augmented vector autoregression (FAVAR) methodology in the spirit of Bernanke, Boivin, and Eliasz [6]. We aggregate the information in a large number of disaggregate real and financial time series into a small number of factors and employ these to evaluate the responses of key time series to a variety of monetary policy instruments. We work in a Bayesian framework and apply Markov Chain Monte Carlo (MCMC) methods to obtain the posteriors. Employing the generalized sign restriction approach toward identification of Amir Ahmadi and Uhlig [2], we find the effects of monetary policy shocks to have been moderate. We invoke Granger causality of monetary instruments, implemented by Bayesian forecasting techniques, to identify the systematic component of monetary policy. Furthermore we analyze the impact of the systematic component of monetary policy by identifying the reaction of the policy instruments to aggregate supply and demand shocks. This broadly confirms the Friedman/Schwartz view that systematic monetary policy was restrictive in late 1929 and again in 1931. However, the effects were again quite moderate. Our results caution against a predominantly monetary interpretation of the Great Depression.

3.1 Introduction

Beginning with the seminal contribution of Friedman and Schwartz [15], the Great Depression has traditionally been associated with restrictive monetary policy. In 1928, the Federal Reserve Bank of New York, then responded to the stock market boom with interest rate hikes from 3.5 % in January to 5 % in September. Between July and October of 1929, it raised its discount rate by another percentage point. After the October stock market crash, the discount rate was reduced again, down to 2 % in January 1931. However, given the rapid decline in price levels, ex-post real interest rates remained high. Monetary authorities also failed to intervene in the banking crisis that unfolded beginning in December of 1930, and interest rates increased again after Britain's departure from the Gold Standard in October, 1931.

This paper is about submitting the role of monetary policy in the Great Depression hypothesis to empirical test. This task is a complex one, as several different channels of monetary policy transmission during the depression have been proposed. The strongest form of the paradigm, expounded by Schwartz, states that both the initial recessionary impulse and the later deepening of the recession were largely caused by the Federal Reserve. In contrast, the original position of Friedman and Schwartz [15] centered more strongly on the role of monetary policy in deepening the slump. This weaker version of the monetary paradigm is also consistent with the emphasis placed on bank panics by Bernanke and others. Bernanke's research stressed financial channels of monetary policy transmission, emphasizing the role of information asymmetries and participation constraints in debtor/creditor relations, as well as of debt deflation. Bernanke and Carey [5] also looked at nominal wage stickiness as an alternative mechanism of monetary policy transmission during the depression.

In the light of these various proposed transmission mechanisms, traditional VAR analysis soon reaches its limits, as it only allows for a small number of time series to model the pertinent dynamics of the money/income causation. One alternative that has been pursued in the recent literature was to obtain counterfactuals from well-specified DSGE models of the Great Depression that focus on one specific monetary transmission mechanism. Bordo, Erceg, and Evans specify a DSGE model with sticky wages, finding evidence in favor of a nominal wage rigidity channel of monetary policy transmission. Christiano, Motto and Rostagno propose a DSGE model with a permanent increase in liquidity preference during the depression, and argue that given this preference shift, easy monetary policy a la Friedman and Schwartz would have mitigated most of the slump.

However, non-monetary interpretations using DSGE techniques seem to have worked equally well in modeling the interwar depression. Prominently, Cole and Ohanian specified a model of collective wage bargaining to argue that real wage rigidity under the New Deal prevented a more complete recovery from the depression after 1933. In a model of international business cycle transmission in the Great Depression, Cole, Ohanian, and Leung examined monetary policy and productivity shocks alongside each other, and found only a minor role for monetary shocks in explaining the slump. On the other extreme of the spectrum of non-monetary models, Harrison and Weder calibrate a sunspot model of investor behavior, which finds strong evidence for an investment-led downturn that was unrelated to monetary policy. Hence, existing research offers a whole menu of interpretations which all seem consistent with the data, although they partly exclude each other.

This is what motivates the approach taken in the present paper. Compared to existing research on the Great Depression, we aim to impose less structure and at the same time analyze a richer dataset. We start out from parsimonious yet informative prior assumptions on the effects of monetary policy. We gear our estimation toward exploiting the information on the common components of business cycle movements in a large cross section of time series. To this end, we employ the factor-augmented vector autoregression (FAVAR) techniques introduced into monetary policy by, among others, Bernanke and Boivin [4] and Bernanke, Boivin, and Eliasz [6] (henceforth BBE). The idea behind this can be interpreted as augmenting the information content in a VAR by a two-step procedure. In a first step, the common dynamics in a large panel of time series are identified using dynamic factor model (DFM) techniques as developed by Geweke [19] and Sargent and Sims [23]. In a second step, the causality between a properly chosen policy instrument and some representative measure of economic activity is examined in a traditional VAR, including the factors as the relevant description of the underlying economic dynamics. Estimation is either in two steps, employing principal-component techniques for DFM part and Maximum Likelihood for the FAVAR, or simultaneous by Bayesian likelihood methods or suitable numerical approximations. In the present paper, we adhere to the Bayesian approach, which allows us to exploit the information on the observables in the VAR specification more completely.

Both traditional VAR analysis and FAVARs for U.S. data have obtained significant but quantitatively small effects of monetary policy on output. In a long-term study on the U.S. since the 1930s, Sims [24] finds that monetary policy on average explains around 12 % of forecast error variance in output. Using the FAVAR technique, Amir Ahmadi and Uhlig [2] report a variance explanation of less than 14 % for industrial output and roughly 10 % for unemployment, order flows, and capacity utilization, evaluated at a 48-months horizon.

VAR evidence on the Great Depression is sparse. Ritschl and Woitek employ time-varying techniques on four different specifications of the monetary transmission mechanism and find that monetary policy explains less than 5 % of output forecast error variance. They also find the forecasting performance of their VARs to be poor. This suggests that a traditional monetary policy VAR, run with the imperfect aggregate data available for the interwar period, might not be able to capture the business cycle dynamics of the Great Depression very well. Given the limitations to data quality in the interwar period, working in a FAVAR framework thus seems particularly promising, as the underlying DFM aggregates information included in a large panel of disaggregate time series. The statistical aggregation procedure implicit in the FAVAR presents an alternative to historical monetary statistics and reconstructed national accounts with their unavoidable interpolations and inaccuracies.

The aim of the present paper is to track the effects of U.S. monetary policy during the interwar years in the data-rich environment provided by the FAVAR approach, and to evaluate them against the postwar evidence collected in previous studies. The Friedman/Schwartz hypothesis on the monetary causes of the Great Depression would suggest that the effects of interwar monetary policy were significant and certainly larger than the rather modest estimates obtained for postwar U.S. data. Any findings that suggest only minor effects of monetary policy would then have to be interpreted as cautioning against a primarily monetary explanation of the Great Depression.

The task at hand is a double one. On the one hand, we follow the standard approach to policy analysis in a VAR, calculating impulse response sequences and forecast error variance decompositions under identifying assumptions about the correlation structure of the VAR residuals. The implicit assumption behind this approach is the neutrality of anticipated monetary policy changes, i.e. of movements along the central bank's reaction function. On the other, we also attempt to trace possible systematic effects of monetary policy, which would be present under a wider set of frictions that allow for (however short-lived) deviations from non-neutrality of movements along the monetary policy reaction function itself. In a VAR, such systematic effects would be identified through Granger causality of monetary policy for other variables of interest. We implement this by taking Bayesian forecasts of key economic indicators from FAVARs with and without past realizations of the policy instrument. Improvements of the forecast in the presence of the policy instrument relative to the baseline would then be an indication of possible systematic policy effects, while the sign of the correction would indicate the expansionary or recessionary stance of systematic policy. This is the closest we can get to providing a test of monetary policy effects in the spirit of Friedman and Schwartz hypothesis. Furthermore we identify the reaction of different policy instruments to aggregate supply and demand shocks tracing the systematic reaction of the monetary authority to changes in the economy.

We proceed in several steps. Section (3.2) presents the basic econometric framework, which closely follows the Bayesian version of the BBE's FAVAR model. Section (3.3) describes the estimation procedure and Section (3.4) presents empirical results for policy shocks from the generalized sign restriction identification approach described in Amir Ahmadi and Uhlig [2]. Section (3.5) and (3.6) looks at possible systematic effects of monetary policy, identifying the policy reaction to aggregate supply and demand shocks, examining Bayesian predictions of key time series with and without the policy instrument at critical junctures. Section (3.7) provides results for the stochastic component of monetary policy and Section (3.8) concludes.

3.2 The Model

The key idea behind dynamic factor models is to parsimoniously represent the comovements in a large set of cross-sectional data by only a limited number of unobserved latent factors[1], which apparently are sufficient to represent the crucial dynamics, and an idiosyncratic component that reflects the variable specific part. The common shocks are referred to the common component, which consists of the factors and the respective factor loadings. In the recent few years there has been a surge in the research on dynamic factor models (henceforth DFM) where advances and several extensions have been introduced. It has become an important tool in empirical macroeconomics since it provides a possibility to break down the dimensionality of the large amounts of economic time series, which have become available and are part of the decision making agents information set, into a few common factors.

3.2.1 Factor-Augmented VAR.

The FAVAR model is very closely related to DFMs and grounds on the idea to combine the standard structural VAR analysis with the recent advances in dynamic factor models, estimating a joint VAR for some variables of interest f_t^y and factors f_t^c extracted from a large panel of informational time series X_t^c. The working hypothesis of the FAVAR model is that while a narrow set of variables f_t^y, notably the policy instrument of the central bank, are perfectly observable and have pervasive effects on the economy, the underlying dynamics of the economy are less perfectly observable, and hence a VAR in just a few key variables would potentially suffer from omitted variable bias. As increasing the size of a VAR is impractical due to problems of dimensionality, the FAVAR approach aims to extract the common dynamics from a wide set of informational indicator series X_t^c, and to include these in the VAR, represented by a small number of factors f_t^c. This approach is well suited for structural analysis such as impulse response analysis and variance decomposition (in particular for the problem at hand). For the estimation procedure the model has to be cast into a state-space representation. The dynamics of the informational variables X_t^c included in the observation

[1] Reichlin et. al. refer to the common components as shocks explaining the same concept through different estimation strategies. The term shock in their case should not be confused with economic shock as e.g. in VARs.

equation are assumed to be driven by observable variables with pervasive effects on the economy (e.g. the central bank's policy instrument), f_t^y, and a small number of unobservable common factors, f_t^c, which together represent the main "driving forces" of the economy, and an idiosyncratic component e_t^c, i.e.

$$X_t^c = \lambda^c f_t^c + \lambda^y f_t^y + e_t^c \tag{3.2.1}$$

$$e_t^c \sim N(0, R_e) \tag{3.2.2}$$

Here λ^f and λ^y denote the matrix of factor loadings of the factors and the policy instrument with dimension $[N_c \times K_c]$ and $[N_c \times N_y]$ respectively. The error term e_t^c has mean 0 and covariance R which is diagonal. Hence the error terms of the observable variables are mutually uncorrelated.

The FAVAR state equation represents the joint dynamics of factors and the observable policy variables (f_t^c, f_t^y).

$$\begin{bmatrix} f_t^c \\ f_t^y \end{bmatrix} = b(L) \begin{bmatrix} f_{t-1}^c \\ f_{t-1}^y \end{bmatrix} + A\nu_t^f \tag{3.2.3}$$

$$u_t^f = A\nu_t \tag{3.2.4}$$

$$u_t^f \sim N(0, Q_u) \tag{3.2.5}$$

where u_t^f is the time t reduced form shock, ν_t^f the time t structural where the contemporaneous relations are represented through the matrix A. The dimensions are $[K \times 1]$, $[K \times 1]$ and $[K \times K]$ respectively where $K = K_c + N_y$ denotes the total number of factors including the perfectly observables ones. In the subsequent estimation we consider the following finite order VAR(P) approximation of the unobserved state dynamics

$$\begin{bmatrix} f_t^c \\ f_t^y \end{bmatrix} = \sum_{p=1}^{P} b_p \begin{bmatrix} f_{t-1}^c \\ f_{t-1}^y \end{bmatrix} + A\nu_t. \tag{3.2.6}$$

3.2.2 Factor Identification.

In order to uniquely identify the model against rotational indeterminacy we have to impose a normalization and additional restrictions. Note that the above state space model could be combined with any orthonormal rotation matrix to which the likelihood is invariant. However any rotation yielding the same likelihood would represent a different representation of the data. Here follow the approach of Bernanke, Boivin and Eliasz [2005] and normalize the upper $[K_c \times K_c]$ block of λ^f to identity and restrict the upper $[K_c \times N_y]$ block of λ^y only contain zeros. Note that this approach is overidentified[2], however, this approach is easy to implement and does not require further restriction on the sign of loadings or further normalizations of the state residual

[2] Alternative restrictions and normalization for the factor identification are reported e.g. in Geweke and Zhou. The analysis and comparison of the different approaches to factor identification goes beyond the scope of this paper.

covariance matrix or the observation residual covariance matrix.

Consequently, Bayesian analysis treats the parameters of the model as random variables. We are interested in inference on the parameter space $\theta = \left(\lambda^f, \lambda^y, R_e, b, Q_u\right)$. Multi move Gibbs Sampling alternately samples the parameters θ and the factors $f_t = \left(f_t^{c\prime}, f_t^{y\prime}\right)^\prime$, given the data. We use the multi move version of the Gibbs sampler (see Carter and Kohn [10] and Frühwirth-Schnatter [16]) because this approach allows us as, a first step, to estimate the unobserved common components, namely the factors via the Kalman filtering technique conditional on the given hyperparameters, and as a second step calculate the hyperparameters of the model given the factors via the Gibbs sampler in the respective blocking.

3.3 Estimation and Identification of Shocks

3.3.1 Estimation

We cast the state space model of the previous section into a stacked first order Markov state space representation. Estimation is facilitated via a multi-move Gibbs sampler which involves the Kalman smoother for evaluating the likelihood of the unobserved factors. Given the sequence of sampled factors we draw the parameters via posterior sampling. In particular we employ a Gibbs sampler for the two blocks of parameters, the first referring to the parameters of the observation equation and the second block contains the parameter space of the state equation. The above state space representation can be rewritten as

$$X_t = \lambda f_t + e_t \tag{3.3.1}$$

$$f_t = \sum_{p=1}^{P} b_p f_{t-p} + u_t^f \tag{3.3.2}$$

where

$$\lambda = \begin{bmatrix} \lambda^f & \lambda^y \\ 0_{N_y \times K_c} & I_{N_y} \end{bmatrix}, \quad R = \begin{bmatrix} R_e & 0 \\ 0 & 0 \end{bmatrix} \tag{3.3.3}$$

where $X_t = \left(X_t^{c\prime}, f_t^{y\prime}\right)^\prime$, $e_t^\prime = \left(e_t^{c\prime}, 0\right)^\prime$ and $f_t = \left(f_t^{c\prime}, f_t^{y\prime}\right)^\prime$. For the companion form of the model we define $F_t = \left({f_t}^\prime, {f_{t-1}}^\prime, ..., {f_{t-p+1}}^\prime\right)^\prime$, $u_t = \left(u_t^{f^\prime}, 0, ..., 0\right)^\prime$, $b = (b_1, b_2, \ldots, b_P)$ and

$$B = \begin{bmatrix} b_1 & b_2 & \ldots & b_{p-1} & b_P \\ I_K & 0 & \ldots & 0 & 0 \\ 0 & I_K & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & I_K & 0 \end{bmatrix}, \qquad Q = \begin{bmatrix} Q_u & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 \end{bmatrix}.$$

where the 0's and Q have dimension $[K \times K]$, and $[PK \times PK]$ respectively. We define $\Lambda \equiv [\lambda\, 0 \ldots 0]$. The final companion form results in

$$F_t = BF_{t-1} + u_t \tag{3.3.4}$$

$$X_t = \Lambda F_t + e_t \tag{3.3.5}$$

The parameter space to be estimated is given by $\theta = (\lambda^y, \lambda^f, b, R_e, Q_u)$ and the history of the observed data and the latent factors is given by $X^T = (X_1, \ldots, X_T)$ and $F^T = (F_1, \ldots, F_T)$ respectively. Hence the estimation algorithm can be simplified and summarized by two steps relying on the blocking scheme. First we initialize the sampler by finding starting values $\theta^0 = (\lambda^{f0}, \lambda^{y0}, b^0, R_e^0, Q_u^0)$ and (F^0). Given a set the initial values (θ^0, F^0) we sample the parameters conditional on the data, and afterwards sample the latent factors given the new set of parameters and data.

Step 1: $F^{T(g)} = p(F^T \mid X^T, \theta^{(g-1)})$

Step 2: $\theta^{(g)} = p(\theta \mid X^T, F^{T(g)})$

We cycle through this procedure sufficient many times until the target distribution has been empirically approximated. An initial number of draws will be discarded, referred to as the burn in and in order to reduce the dependency of the posterior sampler and to reduce the autocorrelation of the chain a thinning parameter $\kappa \geq 1$ can be introduced. Hence only every κ draw after the burn in is stored. The details about the implementation and Specification are reported in section (3.4) on the empirical application. The Algorithm (3.1) contains a pseudo code of the employed algorithm for illustrative purposes. A detailed technical derivation and description of the posterior sampling technique can be found in appendix.

3.3.2 Identification Shocks

One objective of this paper is to analyze the role of monetary policy shocks during the interwar US Great Depression which involves the identification of the non-systematic part of monetary policy via impulse response functions. In order to answer the question raised we have to make identifying assumptions. In this section we will describe the identification strategy that we employ in this paper in order to trace out the impact of the deflationary policy of the central monetary authority during the US Great Depression. Furthermore we are interested in the impact of the systematic part of monetary policy. One way to measure the systematic reaction of the monetary authority is to measure the impact of the policy instrument to other shocks such as aggregate demand and supply shocks. The usual approach to identify the dynamic effects of a shock to monetary policy in our model framework is based on the Cholesky identification where the policy instrument is ordered last in the FAVAR equation after the extracted factors and the other variables of the core VAR such as output and CPI inflation. Thus the policy instrument can react contemporaneously on the other variables through the common factors but not vice versa. As shown in Amir Ahmadi and Uhlig [2] this approach is flawed and produces unreasonable results for the post-war US data. Therefore we will apply sign restrictions which was generalized for the factor model approach in Amir Ahmadi and Uhlig [2] relying on ideas in Uhlig [31]. This

Table 3.1: **Algorithm 1**

Algorithm 1: FAVAR estimation via Multi-move Gibbs sampling.

Step 0, [Initialization]: $p_0(F^0, \lambda^{f0}, \lambda^{y0}, b^0, R_e^0, Q_u^0)$.

```
Set  g ⇝ 0.
Get initial values for states and parameters.
   Set  g ⇝ 1.
```

Step 1, [Evaluate likelihood of latent states]: $p(F^T \mid X^T, \lambda, R_e, b, Q_u) \sim FFBS$

```
   Do forward filtering and backward sampling
```

Step 2, [Sample parameters from observation equation]: $p(\lambda_n, R_{e,nn} \mid X^T, F^T)$

$2.a: \quad p(R_{e,nn} \mid \lambda_n^{(g-1)}, X^T, F^T) \sim f_{\mathcal{IG}}$

$2.b: \quad p(\lambda_n \mid R_{e,nn}^{(g)}, X^T, F^T) \sim f_{\mathcal{N}}$

```
   Sample equation by equation (conditional Gaussianity).
```

Step 3, [Sample parameters from state equation]: $p(b, Q_u \mid X^T, F^T)$

$3.a: \quad p(Q_u \mid F^{T(g)}, X^T) \sim f_{\mathcal{IW}}$

$3.b: \quad p(b \mid Q_u^{(g)}, F^{T(g)}, X^T) \sim f_{\mathcal{N}}$

```
   Sample from normal inverted Wishart density.
   If g ≤ G set g ⇝ g+1 and go to Step 1.
   Otherwise stop.
```

strategy applied to the dynamic factor model framework additionally provides the important possibility to impose restrictions on the impulse response functions on many more variables than usually considered in a standard VAR approach. For further details please refer to the papers listed before.

3.3.2.1 Recursive Identification

The recursive identification scheme as in Christiano, Eichenbaum and Evans imposes a recursive structure of the contemporaneous Matrix relating the reduced form factor residuals into economically meaningful structural shocks. The policy instrument is ordered last, after the common factors and other perfectly observable factors included in the state equation. Let

$$\begin{aligned} Q_u &= E[u_t^f u_t^{f'}] \\ E[u_t^f u_t^{f'}] &= AE[v_t^f v_t^{f'}]A' \\ AE[v_t^f v_t^{f'}]A' &= AA' \end{aligned}$$

where A is lower triangular Cholesky factor of the factor residual covariance matrix and where the policy instrument is ordered last in the FAVAR equation endogenizing monetary policy reaction. The identifying restrictions follow in parts from the factor identification which we employ as explained in Bernanke, Boivin and Eliasz [2005] which is actually independent from the identification task regarding structural shocks. The factor identification restrict the contemporaneous channel through which monetary policy shocks can effect the first $[K_c \times K_c]$ variables of our data panel which is part of the shock identification as the first variables are output related ones. Hence the contemporaneous reaction of the first $[K_c \times K_c]$ output related variables is restricted such that it does not react in the shock period. Furthermore the ordering of the panel is important. A detailed description of the identification assumptions can be found in Stock and Watson [28] and Amir Ahmadi and Uhlig [2].

3.3.2.2 Sign Restriction

Identification of structural shocks through imposing sign restrictions is based on the assumptions about the sign of the impulse response functions of key macroeconomic variables. Such restrictions should represent "'conventional wisdom"' derived from economic theory that most researchers can agree on. In the context of a contractionary monetary policy shocks Uhlig [31] set the restriction prices and nonborrowed reserves should not increase and the federal funds rate should increase for a specified period following the shock. This the approach we employ in this paper.

The structural FAVAR can be arrived at when we premultiply the reduced form version in (3.3.2) with the rotation matrix A, which results in:

$$f_t = \sum_{p=1}^{P} b_p f_{t-p} + A\nu_t^f$$

The crucial step is to represent the one-step ahead prediction error ${\nu^f}_t$ as a linear combination of orthogonalized structural shocks[3]. The fundamental innovations are mutually independent and normalized to have unit variance. Hence $E[{\nu^f}_t {\nu^f}_t{}'] = I_K$. The restriction on A emerges from the covariance structure of the reduced form factor innovation which results in:

$$\begin{aligned} Q_u =& E[u_t^f u_t^{f'}] \\ E[u_t^f u_t^{f'}] =& AE[\nu_t^f \nu_t^{f'}]A' \\ AE[\nu_t^f \nu_t^{f'}]A' =& AA' \end{aligned}$$

According to Uhlig [31] we define an impulse vector as a column of the matrix A. Such a vector can be obtained from any decomposition, e.g. the Cholesky decomposition, of the VCV matrix of the factor residual matrix $\tilde{A}\tilde{A}' = Q_u$.

Definition 1 *The vector $a \in \Re^K$ is called an **impulse vector**, iff there is some matrix A, so that $AA' = Q_u$ and so that a is a column of A*

According to the Proposition 1 of Uhlig [31], any impulse vector can be characterized

[3] See Uhlig [2005]

as follows. Let $\tilde{A}\tilde{A}' = Q_u$ be the Cholesky decomposition. Then a is an impulse vector if and only if there is some K-dimensional vector α of unit length so that

$$a = \tilde{A}\alpha$$

Given the impulse vector, let $r_k(s) \in \Re^K$ be the vector response at horizon s to the k-th shock in a Cholesky-decomposition of Q_u. Then the impulse response $r_a(s)$ for a is simply given by

$$r_a(s) = \sum_{i=1}^{K} \alpha_i r_i(s).$$

For estimation consider the companion form of the state space in (3.3.4)-(3.3.5)

$$\begin{aligned} F_t &= BF_{t-1} + u_t \\ X_t &= \Lambda F_t + e_t \end{aligned}$$

To compute the impulse response vector a, let $\mathbf{a} = [a', 0_{1,K(P-1)}]'$ and compute

$$r_{a,k}(s) = (B^s \mathbf{a})_k.$$

to get the impulse response of factor k to an impulse in a at horizon s. Note that $r_a(s)$ is the vector of impulse response functions of all factors to an impulse vector a at horizon s. As a second step we have to compute the impulse response functions of the single variables by combining the respective factor loading with $r_a(s)$ accordingly. This requires to compute

$$r_a^n(s) = \Lambda_n r_a(s).$$

where Λ_n is the respective n-th row vector of the factor loading matrix. We set the sign restriction on the shape of the individual impulse response functions according to the following assumption:

Assumption 1 A (contractionary) monetary policy impulse vector *is an impulse vector $\mathbf{a}_{\mathbf{money}}$ so that the individual impulse response functions to $\mathbf{a}_{\mathbf{money}}$ of price and nonborrowed reserves are not positive and the impulse responses for the policy instrument such as the short term discount rate (controllable monetary aggregate, e.g. M1) is not negative (positive), for a specified horizon s=0,...,S.*

Assumption 2 A (negative) aggregate demand shock impulse vector *is an impulse vector $\mathbf{a}_{\mathbf{demand}}$ so that the individual impulse response functions to $\mathbf{a}_{\mathbf{demand}}$ of price and output are negative, for a specified horizon s=0,...,S.*

Assumption 3 A (positive) aggregate supply shock impulse vector *is an impulse vector $\mathbf{a}_{\mathbf{supply}}$ so that the individual impulse response functions to $\mathbf{a}_{\mathbf{supply}}$ of price are positive and of output are negative, for a specified horizon s=0,...,S.*

3.4 Empirical Results

3.4.1 Data and Model Specification

All data are taken from the NBER's macroeconomic history database. Most of these data are contemporary and were collected for the business cycle dating project of Burns and Mitchell. Our dataset includes a total of 164 time series, ranging from industrial production to order flows and housing startups, agricultural, raw material, and finished goods prices, measures of deposits, savings, and liquidity in the banking system, as well as interest rates on call money, commercial paper, and various medium and long term bonds. The appendix provides the details along with the NBER macroeconomic database classification codes. To achieve stationarity, some of the data series were transformed into logs, first difference accordingly.

We estimate the model using the data in monthly frequency for the US covering the time span from 1919:02 until 1939:02. This period covers the period of hyperinflation and the downturn of the Great Depression. In the following, we report the results from a FAVAR model with $K_c = 4$ factors and $P = 12$ lags[4] on a dataset including one policy instrument and $N = 164$ informational variables.[5].

3.4.2 Model Fit

We want to assure that our methodology can represent the data in an adequate manner. In this section we want to report results that pursue to assess the models fit to the data. The first obvious check is to see how well the factors represent our panel of data series. To this end, we estimated the R^2s from regressing the respective series onto the five factors. Results are listed in the Table (3.2) below. As can be seen, the overall fit is high; the factors do seem to capture the variance in the individual series very well. Hence, the factor model is informative in the sense that it describes the common components of the interwar business cycle that we are interested in. Thus, a VAR in these factors or common components should not suffer from omitted variable bias, which implies that adding individual series to the VAR in eq. (3) above does not alter the shape of any impulse response functions substantially[6]. As already described above we also estimated the respective marginal contribution of each factor for the explanatory part in the variation of the data. The high value of the R^2 for most variables of interest delivers a first sign of a good representation of the cross section of the data through the factors.

[4] We tried several versions with different lag length (up to 13), but the results compared to the ones reported here do not differ much

[5] We experimented with changing the number of factors, and found little information was added by increasing the dimension of the system by analyzing the marginal contribution of further factor for the explanation of the variation in the data. This is broadly consistent with the results in Stock and Watson [2005], who report an optimal choice of seven factors for their postwar U.S. data set of 132 series with this methodology.

[6] If this property holds strictly, the factor model is termed exact. If including individual series adds to the information content significantly but with small coefficients, the factor model is approximate. See Stock and Watson [2005] for a survey of the implications and for testing strategies.

3.4.3 Explained Variation by Factors.

The model fit is informally measured by the proportion of the variation in the observed is explained the extracted common shocks. From the table (3.2) reporting these results one see that the overall explanation is rather high indicating that the factor structure and the rich interrelations of the different sectors are well captured. The results reported are based on 4 extracted factors. We experimented with extracting up to 7 factors. Te degree of explanation did not change much and in particular the results did not virtually therefore we rely in the subsequent on 4 extracted factors.

Table 3.2: **Estimated R^2 from individual series on FAVAR (DR model).**

Description	R^2	**Description**	R^2
PR IMNF	1	Production (durable mnfct)	0.71
CPI	1	Industrial Production/Trade	0.69
DR	1	Industrial activity	0.68
Total rates charged	1	Business activity growth	0.67
Bankers rates (Customer loans)	1	Index of WSP:	0.61
Open market rates	0.99	WSP: Foods	0.6
CommPR	0.99	General price level	0.56
Yield:Corporate bonds	0.99	Employment: Machinery	0.54
Yields: Corporate bonds	0.99	CPI less food	0.53
Rates on custom. Loans	0.99	PR IPTG	0.53
Rates on custom. Loans (SW)	0.99	Pig Iron	0.51
90day time to money	0.99	Employment: Manufacturing	0.51
Rates on custom. Loans (NW)	0.99	Business activity pittsburgh	0.5
Yields: Public utility	0.98	Index manufacturing prod.	0.5
Banker s accept. Rate	0.98	Steel ingot	0.5
DR Dallas	0.98	Cost of Living index	0.49
DR SF	0.97	Payrolls wkly: Manufacturing	0.49
DR Minneapolis	0.96	Factory payrolls: Machinery	0.49
Yields: Industrial bonds	0.96	Factory payrolls: steel	0.47
Call money rate	0.95	Employment: Steel	0.47
Yield: Long-term bonds	0.95	PR IPDCG	0.45
Yields: Railroad bonds	0.95	WSP Industrial (sensitive Raw)	0.44
Dividend yields	0.95	WSP: Textiles	0.44
Yields: Munipal bonds	0.92	WSP: farm products	0.44
Interest rates FED bank loans	0.91	PR IPCGLA	0.4
PR IPRGD	0.88	Employment: Paper	0.39
WSP: food	0.87	WSP Industrial coomodities	0.39
PR IPDG	0.81	Employment: Lumber	0.37
Yield:Corporate bonds (LG)	0.78	WSP: Building material	0.37
Index business activity	0.77	Employment: Steel	0.37

Here we see the variance decomposition of the factors through the estimated R^2s for each indicator series based on 4 extracted factors.

3.4.4 Convergence Diagnostics

In order to assure that the results are based on converged simulation chains that represent the respective target distribution and e.g. not only some local mode we have to apply further convergence diagnostics for the simulated parameters based on the Gibbs sampler. The respective diagnostics are not a guarantee for convergence but a battery of different convergence diagnostics can at least reduce the uncertainty to some degree. Therefore we employed some diagnostics that are widespread in the MCMC literature which are only briefly explained here and reported in Appendix D.

All our results are based on 50000 simulation draws of the Gibbs sampler of which the first 30000 were discarded as a burn-in to avoid an initial transient of the starting values initializing the simulation. Furthermore we chose a thinning parameter of 5 so that only every fifth draw was kept after the burn-in in order to reduce potential autocorrelation of the sequence. Furthermore we checked the precision of the sampler by plotting the associated error bands.

We furthermore checked the convergence by monitoring the sampler visually through trace plots which shows the evolvement of the sampler and helps to check whether there are jumps in the level of the respective parameter. Furthermore we plotted the first half of the kept draws against the second to check whether the sampler is still *moving* or whether the whole density of the distribution is represented. Few deviation indicates convergence and vice versa

3.5 Forecasting and Granger Causality

This section examines the possible systematic component of monetary policy. Under rational expectations and a minimal set of frictions, as is standard in the surprise Phillips curve paradigm, systematic monetary policy in the sense of movements along the central bank's policy reaction function should be neutral and have no real effects. In this section, we examine the impact of possible non-neutrality of monetary policy, which might be consistent with different patterns of expectation formation, but also with limited use of information, e.g. with rational inattention.

For systematic policy to have real effects, its predictable component should be Granger-causal for key economic variables. We exploit this property by predicting economic series of interest from the monetary policy reaction function, which we implement through a Bayesian FAVAR prediction including the policy instrument. If this forecast is an improvement on an otherwise similar forecast from the same FAVAR without the policy instrument, non-neutrality of the policy instrument may be suspected. At the same time, if the improved forecast including the monetary policy instrument is more recessionary, the likely impact of the systematic policy component appears recessionary, and vice versa.

Friedman and Schwartz [15] provide an extensive discussion of how policy imperfections and incomplete theoretical insights might have induced monetary policy actions that would seem irrational from a modern viewpoint. We do not purport to reverse engineer the revealed preferences of the Federal Reserve but just limit ourselves to tracing their possible consequences.

As in the previous section, we adopt an agnostic stance as to what the actual policy instruments of the Federal Reserve may have been, and obtain results for the discount rate and three different monetary aggregates. We obtain forecasts with and without these policy instruments at five critical junctures during the period. The first includes the information in the FAVAR as of September 1929, the last month before the New York stock market crash. The second includes all data until November 1930, the last month before the first wave of banking panics. The third extends to June 1931, just before the German debt and reparations moratorium, which brought about the downfall of reparations and triggered Britain's departure from the gold Standard. The fourth extends to August 1931, the last month before Britain indeed broke away from the Gold Standard. The last forecast is based on information up until February 1933, the month before Roosevelt's bank closure and the formal inception of the New Deal.

As of September 1929, the FAVAR is unable to predict the imminent depression: after a brief dip, it forecasts a strong seasonal upturn in the spring of 1930, followed by a very gradual slide into recession. In contrast, the forecast from November 1930 is consistently worse than realized manufacturing activity during 1931. The later forecasts from the non-monetary FAVAR are all roughly consistent with later realizations.

These non-monetary FAVARs appear to bear out conventional wisdom about the early phase of the slump, as laid out in Friedman and Schwartz [15] and Bernanke, or in Temin: the sharp downturn after 1929 was itself not predictable. Interestingly, however, the banking panics do not appear to worsen the outlook further: the non-monetary FAVAR forecast of output predicts doom and gloom for 1931 even before the first banking panic of December 1930.

Predicting output from the same FAVAR, but with the monetary policy instrument included, is tantamount to projecting output from the dynamic monetary policy reaction function. However, not all observations entering into the FAVAR estimate of the reaction function receive equal weight. As the Bayesian forecasts we applied include a geometric shrinkage prior, the most recent of these observations before the forecasting period receive the most weight. This is reassuring as it implies that the forecasts will indeed reflect the stance of systematic monetary policy in the last months before the forecast. Forecasting figures reporting the forecasts obtains these forecasts for a variety of possible monetary policy instruments. From September 1929, all of these forecasts (shown in the uppermost row) predict a substantial increase in output during 1930. This implies that in contrast to conventional wisdom, the monetary stance viewed from 1929 was expansionary. This result hold irrespective of the monetary policy instrument under consideration, and is thus independent of whether or not the monetarist quantity of money paradigm is accepted by the researcher. However, all forecasts from the monetary FAVAR based on September 1929 data perform substantially worse than the non-monetary FAVAR, indicating only a minor role for systematic monetary policy in the unfolding events.

Systematic monetary policy again comes out more expansionary than the non-monetary baseline for late 1930. As all money FAVAR forecasts are closer to the actual output data than the non-monetary FAVAR for 3-5 months, the implication would be that a slight expansionary stance of systematic monetary policy during that period may have had a stabilizing influence.

Policy signals for mid-1931 are contradictory. While the interest instrument suggests an expansionary but ineffective monetary stance, the forecasts from the FAVARs including high powered money and M1 indicate a strong, but again ineffective, recessionary slant. The forecast from M2 is close to the actual trajectory, owing probably to the strong endogenous components in this monetary aggregate that make its use as a policy instrument doubtful. Mixed and, again, ineffective policy positions emerge from the forecasts as of August 1931, although the medium term forecasts from the money FAVARs look more expansionary again. Interest rate policy continues to look contractionary through early 1933, although the monetary aggregates, notably high powered money, indicate an increasingly expansionary stance. This last bit of evidence is reminiscent of the result of Lucas and Rapping, who concluded that prior to the New Deal, monetary policy had changed its stance toward expansion. Again, however, the comparison with the non-monetary FAVAR is sobering: there is hardly any improvement in the forecast upon adding a monetary policy instrument. Except possibly for the first half of 1931, systematic monetary policy does not seem to have had much influence on the course of the depression.

In stark contrast to the negative expectations about outlook implicit in the FAVAR, deflation itself appears to be quite unpredictable: all forecasts from the non-monetary FAVAR invariably predict a rebound in CPI levels (see the forecasting without policy instrument). Only the last of these forecasts, based on data to February 1933, roughly predicts CPI inflation correctly.

This evidence is again consistent with conventional wisdom, see in particular Hamilton (1987, 1992). The interesting bit here is that there is no evidence of learning or updating about the deflationary process; the priors in the forecast of CPI appear impossible to overturn. Taking this further, if the FAVAR aggregates the information available to monetary decision makers at the time, their lack of worries about further monetary easing becomes apparent: given the strong deflationary signals that monetary policy was emitting, no further action seemed necessary or even useful. Monetary policy in the conventional sense had lost traction in 1929, and apparently die not regain it before well into the 1930s.

3.6 The Systematic Component of Monetary Policy

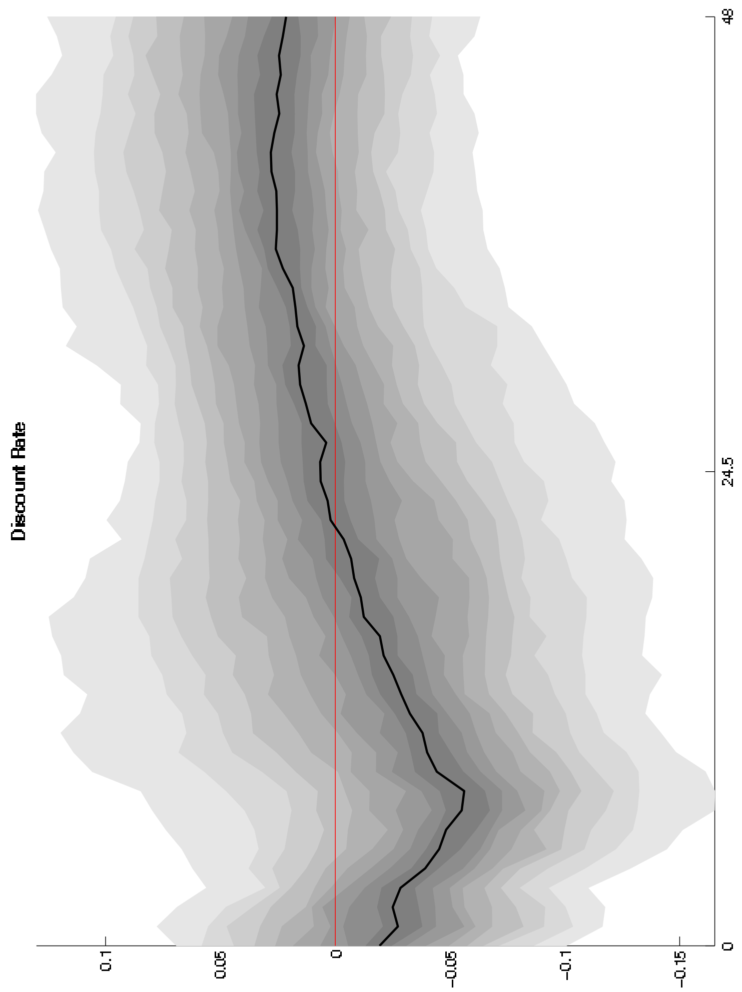
As a second task to recover the impact of systematic monetary policy during the great depression in a more direct way we identify the reaction of the monetary policy instruments to aggregate negative demand shocks and positive supply shocks. We can directly measure if and how the monetary authority reacted to change in output and prices and compare that to the postwar policy regime in order to judge if and to what extend monetary policy was responsive to changes in inflation and output. In the following two subsections we report the results for demand and supply shocks.

3.6.1 Aggregate Supply Shocks

3.6.1.1 Full Sample Analysis

The restrictions we to identify the reaction of the policy instrument following a positive aggregate supply shock is to impose the impulse response of CPI inflation to be negative and the impulse response of FRB index of manufacturing to be positive for a horizon of 6 month. Indeed we find that the policy instruments and hence systematic monetary policy is not responsive to supply shocks increasing the price level and reducing output. This hold true for all 5 model specifications. The monetary aggregates **M1** and **M2** are not responsive at all. Merely the short term interest rates **Commercial Paper Rates** and **Discount Rates** increase following a positive aggregate supply shock in the M2 model. The complete set of results for supply shocks for all subperiods and models can be found in the Appendix.

Figure 3.1: **Impulse Responses to Supply Shock (Full sample): All Models**



3.6.1.2 Subsample Analysis

Turning to the subsample analysis we find that a somewhat different picture emerges for the 5 subsample periods for the **Commercial Paper Rate Model** and the **Discount Rate Model**. The subsamples are again at the five critical junctures during the period reported in section (3.5). The first includes the information in the FAVAR as of September 1929, the last month before the New York stock market crash. The second includes all data until November 1930, the last month before the first wave of banking panics. The third extends to June 1931, just before the German debt and reparations moratorium, which brought about the downfall of reparations and triggered Britain's departure from the gold Standard. The fourth extends to August 1931, the last month before Britain indeed broke away from the Gold Standard. The last is based on information up until February 1933, the month before Roosevelt's bank closure and the formal inception of the New Deal. Here we find an increase in Short term interest rates up to 2 years. Turning to the three monetary aggregate models reaction of the short term interest rates increases for all the subperiods except for the first one.

3.6.2 Aggregate Demand Shocks

Regarding the reaction of the policy instruments following a negative aggregate demand shock we find for the full sample model no reaction of the monetary aggregates. Short term interest rates slightly decrease following a negative demand shock. This holds only for the **Commercial Paper Rate Model** and the **Discount Rate Model** though with a rather high degree of uncertainty.

3.6.2.1 Subsample Analysis

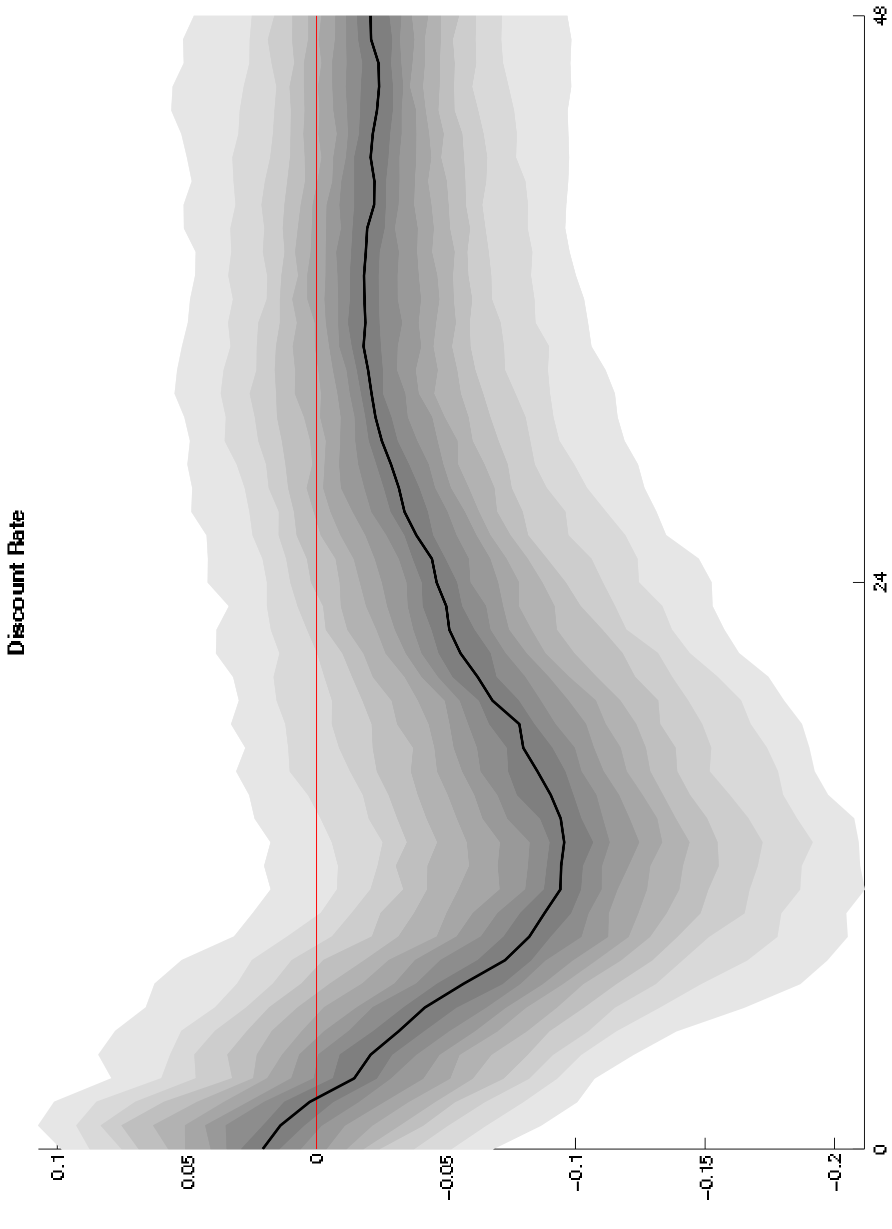
Turning to the subsample analysis we find the same picture for the first to subperiods. There is hardly any reaction of the monetary authority to demand shocks decreasing output and prices. However for the remaining three subperiods short term interest rates clearly decrease following a negative demand shock. This suggest that systematic monetary policy during the Great depression started to react to demand shocks before the last month before the first wave of banking panics. This result holds for all the five models considered. The full list of results for a demand shock are reported in the Appendix.

3.7 Monetary Policy Shock

The resilience of the price puzzle inherent the FAVAR exercises combined with Cholesky identification[7] suggests that, contrary to the hopes expressed in BBE [2005], augmenting the VAR by factors distilled from a large panel of time series may not be sufficient to ensure well behaved impulse response functions. This also implies that even in FAVARs, identification by temporal exclusion restrictions may lead to mismeasurement of the real effects of monetary policy. Consequently, the whole issue of correct identification in VARs emerges again, which motivates us to repeat the exercises

[7]Due to space limitations results are not shown here and can be sent upon request.

Figure 3.2: **Impulse Responses to Demand Shock (Full sample): All Models**



with Uhlig's sign restriction identification. The intuition behind this approach is that in modeling the impulse response functions, the researcher is drawing from realizations of the MA representation of the (FA)VAR based on a diffuse prior, and discards all realizations that are inconsistent with a prior sign restrictions. We implement this by imposing a sign restriction on the response of the CPI to a contractionary monetary shock.

The responses of most real variables are hump- or s-shaped, veering into positive (but mostly insignificant) around a one-year lag and swinging back into negative around the three-year lag, sometimes significantly so. The reactions of the FRB index of manufacturing are near-significant both at the one year and after the three year lag, first with a positive response, then with a negative one.

The variance decompositions show most of the real effect on the indices of business activity after three years, with a peak around 20 %. In contrast, the explained part of the variance of the FRB index of manufacturing remains solidly below 10 %, averaging between 6 and 7 % over the four-year horizon that we look at. This seems close to the values reported by BBE for postwar industrial output.

Repeating the exercise with the commercial paper rate as the policy instrument, the impulse responses (in Figures 13 and 14) all have a characteristic S-shape, reacting positively after one year and swinging back into negative after two years.
However, even at the trough the responses are only near-significant. Near this trough, the variance decompositions show a peak of 10 % of explained forecast error variance for the indices of business cycle activity, and of roughly 8 % for the FRB index of manufacturing. The average and cumulative variance explanations are again below that.

Drawing the results of this section together, we find that the responses of the real economy to contrationary monetary shocks are in generally weak and, pathologically, change their signs. This result obtains under two different specifications of the monetary policy instrument and two different identification schemes for the innovations in the VAR. The evidence for pervasive negative effects of contractionary monetary policy during the Great Depression is not robust.

3.8 Conclusion

Recent research has attempted to increase the explanatory power of vector autoregressions for monetary policy analysis by drawing on the common components in a large panel of time series. In this paper, we employed the factor augmented vector autoregression (FAVAR) methodology of Bernanke, Boivin and Eliasz [2005] to reassess the effects of monetary policy on the U.S. economy during the interwar Great Depression.

We used a panel of 164 time series, taken from the macroeconomic history database of the NBER, to provide information on the common component of the U.S. business cycle during the interwar period. We specified FAVARs based on this information set for two specifications of the monetary policy instrument and two identification schemes for the innovations to the VAR.

We find that while monetary policy was clearly not neutral, its effects on the real economy were mixed and changed signs. Also, we find the overall contribution of monetary policy to the variance explanation of real variables to be as low as in the postwar period, if not lower. We also find that under the traditional temporal ordering of innovations to the VAR, the price puzzle is back, in spite of the large amount of information fed into the dynamic factors in the VAR. Under a sign restriction approach, the results come out slightly sharper but are still far from telling a clear-cut story about the pervasive effects of monetary policy during the Great Depression. At the present stage, we conclude that while monetary policy effects certainly played some role in the interwar depression, there is only scant support for the traditional hypothesis that the Great Depression was mostly a monetary phenomenon.

Appendix B

Appendix B.1: Bayesian Inference based on MCMC

Bayesian analysis treats the parameters of the model as random variables. We are interested in inference on the parameter space $\Theta = \left(\Lambda^f, \Lambda^y, R, vec(\Phi), \Sigma_v\right)$ and the factors $\{F_t\}_{t=1}^{T}$. Multi move Gibbs Sampling alternately samples the parameters θ and the factors F_t, given the data. We use the multi move version of the Gibbs sampler because this approach allows us as, a first step, to estimate the unobserved common components, namely the factors via the Kalman filtering technique conditional on the given hyperparameters, and as a second step calculate the hyperparameters of the model given the factors via the Gibbs sampler in the respective blocking.

Let $\widetilde{X}_T = (X_1, \ldots, X_T)$ and $\widetilde{F}_T = (F_1, \ldots, F_T)$ define the histories of X and F, respectively. The task is to derive the posterior densities which require to empirically approximate the marginal posterior densities of F and Θ:

$$p(\widetilde{F}_T) = \int p(\widetilde{F}_T, \theta) d\Theta$$
$$p(\Theta) = \int p(\widetilde{F}_T, \Theta) d\widetilde{F}_T$$

where

$$p(\widetilde{F}_T, \Theta)$$

is the joint posterior density and the integrals are taken with respect to the supports of Θ and F_T respectively. The procedure applied to obtain the empirical approximation of the posterior distribution is the previously mentioned multi move version of the Gibbs sampling technique by Carter and Kohn [1994] which is also applied by BBE[8].

3.8.1 B.1.1 Choosing the Starting Values Θ^0

In general one can start the iteration cycle with any arbitrary randomly drawn set of parameters, as the joint and marginal empirical distributions of the generated parameters will converge at an exponential rate to its joint and marginal target distributions as $S \to \infty$. This has been shown by Geman and Geman [1984]. We will try several starting values in order to assure that our model has converged and does not depend on the choice of initial values. We follow the advice of Eliasz [2005] that one should judiciously select the starting values in the framework of large dimensional models. In case of large cross-sections, highly dimensional likelihoods make irregularities more likely. This can reduce the number of draws relevant for convergence and hence saves

[8] For more details see Kim and Nelson [1999], Eliasz [2005] and BBE [2005]

time, which in a computer-intensive statistical framework is of great relevance. We apply the first step estimates of principal component analysis to select the starting values. Since Gelman and Rubin [1992] have shown that a single chain of the Gibbs sampler might give a "false sense of security ", it has become common practice to try out different starting values, at best from a randomly (over)dispersed set of parameters, and then check the convergence verifying that they lead to similar empirical distributions.

3.8.2 B.1.2 Conditional density of the factors $\{F_t\}_{t=1}^{T}$ given $\widetilde{X}_T$ and Θ

In this subsection we want to draw from

$$p(\widetilde{F}_T \mid \widetilde{X}_T, \Theta)$$

assuming that the hyperparameters of the parameter space Θ are given, hence we describe Bayesian inference on the dynamic evolution of the factors F_t conditional on X_t for $t = 1, \dots, T$ and conditional on Θ. The transformations that are required to draw the factors have been done in the previous section. The conditional distribution, from which the state vector is generated, can be expressed as the product of conditional distributions by exploiting the Markov property of state space models in the following way

$$p(\widetilde{F}_T \mid \widetilde{X}_T, \Theta) = p(F_T \mid \widetilde{X}_T, \Theta) \prod_{t=1}^{T-1} p_F(F_t \mid F_{t+1}, \widetilde{X}_T, \Theta)$$

The state space model is linear and Gaussian, hence we have:

$$F_T \mid \widetilde{X}_T, \Theta \sim N(F_{T|T}, P_{T|T}) \tag{3.8.1}$$

$$F_t \mid F_{t+1} X_T, \Theta \sim N(F_{t|t,F_{t+1}}, P_{t|t,F_{t+1}}) \tag{3.8.2}$$

with

$$F_{T|T} = E(F_T \mid \widetilde{X}_T, \Theta) \tag{3.8.3}$$

$$P_{T|T} = Cov(F_T \mid \widetilde{X}_T, \Theta) \tag{3.8.4}$$

$$F_{t|t,F_{t+1}} = E(F_t \mid \widetilde{X}_T, F_{t+1}, \Theta) = E(F_t \mid F_{t+1}, F_{t|t}, \Theta) \tag{3.8.5}$$

$$P_{t|t,F_{t+1}} = Cov(F_t \mid \widetilde{X}_T, F_{t+1}, \Theta) = Cov(F_t \mid F_{t+1}, F_{t|t}, \Theta) \tag{3.8.6}$$

where (3.8.1) holds for the Kalman filter for $t = 1, \dots, T$ and (3.8.2) holds for the Kalman smoother for $t = T-1, T-2, \dots, 1$. Here $F_{t|t}$ refers to the expectation of F_t conditional on information dated t or earlier. We can, then, obtain $F_{t|t}$ and $P_{t|t}$ for $t = 1, \dots, T$ by the Kalman Filter, conditional on Θ and the data $\widetilde{X}_T$, by applying the formulas in Hamilton (1994), for example. From the last iteration, we obtain $F_{T|T}$ and $P_{T|T}$ and using those, we can draw F_t. We can go backwards through the sample, deriving $F_{T-1|T-1,F_t}$ and $P_{T-1|T-1,F_t}$ by Kalman Filter, drawing F_{T-1} from (14), and so on for F_t, $t = T-2, T-3, \dots, 1$. A modification of the Kalman filter procedure, as described in Kim and Nelson (1999), is necessary when the number of lags p in the FAVAR equation is greater than 1.

3.8.3 B.1.3 Conditional density of the parameters Θ given $\widetilde{X}_T$ and $\{F_t\}_{t=1}^T$

Drawing from the conditional distribution of the parameters $p(\Theta \mid \widetilde{X}_T, \widetilde{F}_T)$ can be blocked into to parts, namely the one referring to the observation equation and the second part referring to the state equation.

3.8.3.1 B.1.3.1 Conditional density of Λ and R

This part refers to observation equation of the state space model which, conditional on the estimated factors and the data, specifies the distribution of Λ and R. Here we can apply equation by equation OLS in order to obtain $\hat{\Lambda}$ and $\hat{Z}$. This is feasible due to the fact that the errors are uncorrelated. According to the specification by BBE we also assume a proper (conjugate) but diffuse inverse Gamma prior for each σ_n^2:

$$R_{ii}^{prior} \sim \mathcal{IG}(3, 0.001)$$

Note that R is assumed to be diagonal. The posterior then has the following form

$$R_{ii}^{posterior} \mid X_T, F_T \sim \mathcal{IG}(\bar{R}_{ii}, T + 0.001)$$

where $\bar{R}_{ii} = 3 + \hat{Z}_i'\hat{Z}_i + \hat{\Lambda}_i'[M_0^{-1} + (F_T^{(i)'} F_T^{(i)})^{-1}]^{-1}\hat{\Lambda}_i$ and M_0^{-1} denoting the variance parameter in the prior on the coefficients of the i-th equation of Λ_i. The normalization discussed in section (4) in order to identify the factors and the loadings separately requires to set $M_0 = I$. Conditional on the drawn value of R_{ii} the prior on the factor loadings of the i-th equation is:

$$\Lambda_i^{prior} \sim \mathcal{N}(0, R_{ii} M_0^{-1}).$$

The regressors of the i-th equation are represented by $\tilde{\mathbf{F}}_T^{(i)}$. The values of Λ_i are drawn from the posterior

$$\Lambda_i^{posterior} \sim \mathcal{N}(\bar{\Lambda}_i, R_{ii}\bar{M}_i^{-1})$$

where $\bar{\Lambda}_i = \bar{M}_i^{-1}(F_T^{(i)'} F_T^i)\hat{\Lambda}_i$ and $\bar{M}_i^{-1}(F_T^{(i)'} F_T^i)$.

3.8.3.2 B.1.3.2 Conditional density of $vec(\Phi)$ and Σ_ν

The next Gibbs block requires to draw $vec(\Phi)$ and Σ_ν conditional on the most current draws of the factors, the $R_{ii}'s$ and $\Lambda_i's$ and the data. As the FAVAR equation has a standard VAR form one can likewise estimate $vec(\hat{\Phi})$ and $\hat{\Sigma_\nu}$ via equation by equation OLS. We impose a diffuse conjugate Normal-Wishart prior:

$$\begin{aligned} vec(\Phi)^{prior} \mid \Sigma_\nu \sim & \quad \mathcal{N}(0, \Sigma_\nu \otimes \Omega_0) \\ \Sigma_\nu^{prior} \sim & \quad \mathcal{IW}(\Sigma_{\nu,0}, K + M + 2) \end{aligned}$$

which results in the following posterior:

$$vec(\Phi)^{posterior} \sim \mathcal{N}(vec(\bar{\Phi}), \Sigma_\nu \otimes \bar{\Omega})$$
$$\Sigma_\nu^{posterior} \sim \mathcal{IW}(\bar{\Sigma_\nu}, T + K + M + 2)$$

In the spirit of the Minnesota prior, it is desirable to have a prior which assigns less impact to more distant lags. Hence, the BBE [2005] specification follows Kadiyala and Karlsson [1997]. First we draw Σ_ν from the posterior, where $\bar{\Sigma_\nu} = \Sigma_{\nu,0} + \hat{V}'\hat{V} + \hat{\Phi}'[\Omega_0 + (F'_{T-1}F_{T-1})^{-1}]^{-1}\hat{\Phi}$ and where $\hat{V}$ is the matrix of OLS residuals. Then, conditional on the draw Σ_ν we draw from the posterior of the coefficients where $\bar{\Phi} = \bar{\Omega}(F'_{T-1}F_{T-1})\hat{\Phi}$ and $\bar{\Omega} = (\Omega_0^{-1} + (F'_{T-1}F_{T-1}))^{-1}$. To ensure stationarity, we truncate the draws and only accept values for Φ less than one in absolute values . This block on Kalman filter and smoother and the block on drawing the parameter space are iterated until convergence is achieved. For the implementation of the DCNW prior it required to set the diagonal elements of $\Sigma_{\nu,0}$ to the corresponding d-lag univariate autoregressions, σ_i^2. We construct the diagonal elements of Ω_0 such that the prior variances of the parameter of the k lagged j'th variable in the i'th equation equals $\sigma_i^2/k\sigma_j^2$.[9]

Appendix B.2 Data

All data are taken from the NBER's macroeconomic history database. Most of these data are contemporary and were collected for the business cycle dating project of Burns and Mitchell (1947). Our dataset includes a total of 123 time series.

Pos	NBER Code	Description	SA	TC
1	1130	PIG IRON PRODUCTION	YES	2
2	4051	INDEX OF THE GENERAL PRICE LEVEL	YES	2
3	13012	FEDERAL RESERVE BANK DISCOUNT RATES, SAN FRANCISCO	YES	2
4	14125	CURRENCY HELD BY THE PUBLIC	YES	2
5	1054	INDEX OF PRODUCTION OF MANUFACTURES, SEASONALLY ADJUSTED	YES	2
6	1055	INDEX OF PRODUCTION OF PRODUCERS GOODS	YES	2
7	1056	INDEX OF PRODUCTION OF CONSUMERS GOOD	YES	2
8	1057	INDEX OF PRODUCTION OF CONSUMERS GOODS, EXCLUDING AUTOMOBILES	YES	2
9	01057A	INDEX OF PRODUCTION OF DURABLE GOODS	YES	2
10	01057B	INDEX OF PRODUCTION OF TRANSIENT GOODS	YES	2
11	1058	WHEAT FLOUR PRODUCTION	YES	2
12	1060	CORN GRINDINGS	YES	2
13	1064	TOTAL MEAT CONSUMPTION	YES	2
14	1071	BUTTER CONSUMPTION	YES	2
15	1105	PAPER PRODUCTION, ALL GRADES	YES	2
16	01125A	CRUDE PETROLEUM CONSUMPTION, RUNS TO STILLS	YES	2
17	1126	GASOLINE PRODUCTION AT REFINERIES	YES	2
18	1131	MERCHANT PIG IRON PRODUCTION	YES	2
19	1135	STEEL INGOT PRODUCTION	YES	2
20	1144	AUTOMOBILE PRODUCTION, TRUCKS	YES	2
21	1148	RAILROAD LOCOMOTIVE SHIPMENTS, DOMESTIC, BY CAR BUILDERS	YES	2
22	1149	FREIGHT CAR SHIPMENTS, DOMESTIC	YES	2
23	1171	WOODWORKING MACHINERY SHIPMENTS, VALUE	YES	2
24	1175	INDEX OF PRODUCTION OF MANUFACTURES, TOTAL	YES	2
25	01191B	INDEX OF COMMERCIAL PRODUCTION OF FOODSTUFFS AND TOBACCO	YES	2
26	1204	INDEX OF PRODUCTION OF FUELS	YES	2
27	1234	INDEX OF PRODUCTION OF DURABLE MANUFACTURES	YES	2
28	1260	INDEX OF PRODUCTION OF MANUFACTURED FOOD PRODUCTS	YES	2
29	3009	FREIGHT CAR SURPLUS	YES	2
30	03016A	OPERATING REVENUES OF RAILROADS, PASSENGER	YES	2

[9] For a detailed discussion of the implementation of the prior see the NBER working paper version of BBE (2004) and Kadiyala and Karlsson (1997).

Pos	NBER Code	Description	SA	TC
31	03016B	OPERATING REVENUES OF RAILROADS, FREIGHT	YES	2
32	4001	WHOLESALE PRICE OF WHEAT, CHICAGO, SIX MARKETS	YES	2
33	4005	WHOLESALE PRICE OF CORN, CHICAGO	YES	2
34	4006	WHOLESALE PRICE OF COTTON, NEW YORK; 10 MARKETS	YES	2
35	4007	WHOLESALE PRICE OF CATTLE, CHICAGO	YES	2
36	4008	WHOLESALE PRICE OF HOGS, CHICAGO	YES	2
37	4015	WHOLESALE PRICE OF COPPER, ELECTROLYTE, NEW YORK	YES	2
38	4017	WHOLESALE PRICE OF PIG LEAD, NEW YORK	YES	2
39	4030	WHOLESALE PRICE OF GRANULATED SUGAR	YES	2
40	4034	WHOLESALE PRICE OF COFFEE	YES	2
41	4048	INDEX OF WHOLESALE PRICES, BUREAU OF LABOR STATISTICS	YES	2
42	4052	CONSUMER PRICE INDEX, ALL ITEMS LESS FOOD	YES	2
43	4058	INDEX OF WHOLESALE PRICES OF FARM PRODUCTS	YES	2
44	4061	INDEX OF WHOLESALE PRICES OF FOODS	YES	2
45	4064	INDEX OF WHOLESALE PRICE OF TEXTILES	YES	2
46	4066	WHOLESALE PRICES OF METAL AND METAL PRODUCTS	YES	2
47	4068	INDEX OF WHOLESALE PRICES OF BUILDING MATERIALS	YES	2
48	4071	INDEX OF RETAIL PRICES OF FOOD AT HOME	YES	2
49	4074	WHOLESALE PRICE OF OATS, CHICAGO	YES	2
50	4072	COST OF LIVING INDEX	YES	2
51	4079	WHOLESALE PRICE OF CRUDE PETROLEUM, AT WELLS	YES	2
52	4092	WHOLESALE PRICE OF SLAB ZINC	YES	2
53	4099	WHOLESALE PRICE OF COMMON BRICKS, DOMESTIC, NEW YORK	YES	2
54	4128	CONSUMER PRICE INDEX, ALL ITEMS	YES	2
55	4129	WHOLESALE PRICE OF TEA	YES	2
56	4134	WHOLESALE PRICE OF STRUCTURAL STEEL	YES	2
57	4181	WHOLESALE PRICE OF STEEL RAILS	YES	2
58	4189	INDEX OF WHOLESALE PRICES OF INDUSTRIAL COMMODITIES, BABSON	YES	2
59	4202	INDEX OF WHOLESALE PRICES OF 15 SENSITIVE INDUSTRIAL RAW	YES	2
60	06002A	INDEX OF DEPARTMENT STORE SALES	YES	2
61	06002B	THE PHYSICAL VOLUME OF DEPARTMENT STORE SALES	YES	2
62	6008	SALES BY GROCERY CHAIN STORES	YES	2
63	6009	VARIETY CHAIN STORE SALES, ADJUSTED FOR TREND, PRICE	YES	2
64	6029	INDEX OF ORDERS FOR MACHINE TOOLS AND FORGING MACHINERY	YES	2
65	6058	INDEX OF TOTAL ADVERTISING	YES	2
66	6059	INDEX OF WHOLESALE SALES OF SHOES	YES	2
67	7001	DOMESTIC EXPORTS OF CRUDE FOODSTUFFS	YES	2
68	7002	DOMESTIC EXPORTS OF MANUFACTURED FOODSTUFFS	YES	2
69	7004	DOMESTIC EXPORTS OF SEMI-MANUFACTURES	YES	2
70	7005	DOMESTIC EXPORTS OF FINISHED MANUFACTURES	YES	2
71	7012	IMPORTS FOR CONSUMPTION OF CRUDE FOOD STUFFS	YES	2
72	7013	IMPORTS OF MANUFACTURED FOODSTUFFS	YES	2
73	7014	IMPORTS FOR CONSUMPTION OF CRUDE MATERIALS	YES	2
74	7015	IMPORTS FOR CONSUMPTION OF SEMI-MANUFACTURES	YES	2
75	7016	IMPORTS FOR CONSUMPTION OF FINISHED MANUFACTURES	YES	2
76	7023	TOTAL EXPORTS	YES	2
77	7028	TOTAL IMPORTS	YES	2
78	8010B	PRODUCTION WORKER EMPLOYMENT, MANUFACTURING, TOTAL	YES	2
79	8014	INDEX OF FACTORY EMPLOYMENT, PAPER AND PRINTING	YES	2
80	8015	INDEX OF FACTORY EMPLOYMENT, IRON AND STEEL PRODUCTS	YES	2
81	8016	INDEX OF FACTORY EMPLOYMENT, STONE, CLAY AND GLASS PRODUCTS	YES	2
82	8017	INDEX OF FACTORY EMPLOYMENT, LUMBER AND PRODUCTS	YES	2
83	8018	INDEX OF FACTORY EMPLOYMENT, MACHINERY	YES	2
84	8046	AVERAGE WEEKLY EARNINGS, REPRESENTATIVE FACTORIES	YES	2
85	8061	INDEX OF COMPOSITE WAGES	YES	2
86	8069	INDEX OF AGGREGATE WEEKLY PAYROLLS, PRODUCTION WORKERS TOTAL MANUFACTURING	YES	2
87	8071	INDEX OF FACTORY PAYROLLS, TEXTILES	YES	2
88	8072	INDEX OF FACTORY PAYROLLS, PAPER AND PRINTING	YES	2
89	8073	INDEX OF FACTORY PAYROLLS, IRON AND STEEL PRODUCTS	YES	2
90	8074	INDEX OF FACTORY PAYROLLS, STONE CLAY AND GLASS	YES	2
91	8075	INDEX OF FACTORY PAYROLLS - LUMBER AND PRODUCTS	YES	2
92	8076	INDEX OF FACTORY PAYROLLS, MACHINERY	YES	2
93	8078	INDEX OF FACTORY PAYROLLS, NEW YORK STATE FACTORIES	YES	2
94	8088	INDEX OF FACTORY EMPLOYMENT-BAKING	YES	2
95	8101	INDEX OF FACTORY EMPLOYMENT, LEATHER AND MANUFACTURES	YES	2
96	8104	INDEX OF FACTORY EMPLOYMENT, PAPER AND PULP	YES	2
97	8106	INDEX OF EMPLOYMENT, HARDWARE	YES	2
98	8110	INDEX OF FACTORY PAYROLLS, CANE SUGAR REFINING	YES	2
99	8112	INDEX OF FACTORY PAYROLLS, BAKING	YES	2
100	8114	INDEX OF FACTORY PAYROLLS, TOBACCO MANUFACTURES	YES	2

Pos	NBER Code	Description	SA	TC
101	8145	INDEX OF FACTORY PAYROLLS, AUTOMOBILES	YES	2
102	8261	AVERAGE WEEKLY EARNINGS, MANUFACTURING, TOTAL	YES	2
103	11001	BOND SALES, PAR VALUE	YES	2
104	11005	AMERICAN RAILROAD STOCK PRICES	YES	2
105	11009	INDUSTRIAL STOCK PRICE INDEX, DOW-JONES	YES	2
106	11025	INDEX OF ALL COMMON STOCK PRICES, COWLES COMMISSION AND S& P CORPORATION	YES	2
107	12002A	INDEX OF INDUSTRIAL ACTIVITY	YES	2
108	12003	INDEX OF AMERICAN BUSINESS ACTIVITY	YES	2
109	12004	INDEX OF INDUSTRIAL PRODUCTION AND TRADE	YES	2
110	12007	INDEX OF AMERICAN BUSINESS ACTIVITY	YES	2
111	12009A	INDEX OF BUSINESS ACTIVITY, PITTSBURGH	YES	2
112	12009	INDEX OF AGRICULTURAL MARKETINGS	YES	2
113	12013	BANK CLEARINGS, DAILY AVERAGE	YES	2
114	13001	CALL MONEY RATES, MIXED COLLATERAL	YES	2
115	13002	COMMERCIAL PAPER RATES, NEW YORK CITY	YES	2
116	13003	NINETY DAY TIME-MONEY RATES ON STOCK EXCHANGE LOANS	YES	2
117	13004	RATES ON CUSTOMER LOANS, NEW YORK CITY	YES	2
118	13005	RATES ON CUSTOMERS LOANS, NORTHERN AND WESTERN CITIES	YES	2
119	13006	BANK RATES ON CUSTOMERS LOANS, SOUTHERN AND WESTERN CITIES	YES	2
120	13007	BANKER S ACCEPTANCE RATES, NEW YORK CITY	YES	2
121	13008	INTEREST RATES ON FEDERAL LAND BANK LOANS, TWELVE FEDERAL LAND BANKS	YES	2
122	13009	DISCOUNT RATES, FEDERAL RESERVE BANK OF NEW YORK	YES	2
123	13010	FEDERAL RESERVE BANK DISCOUNT RATES, MINNEAPOLIS	YES	2
124	13011	FEDERAL RESERVE BANK DISCOUNT RATE, DALLAS	YES	2
125	13021	INDEX OF YIELDS OF HIGH GRADE CORPORATE AND MUNICIPAL BONDS	YES	2
126	13023	INDEX OF YIELDS OF HIGH GRADE MUNICIPAL BONDS	YES	2
127	13024	YIELDS OF HIGH GRADE RAILROAD BONDS	YES	2
128	13025	INDEX OF YIELDS OF HIGH GRADE PUBLIC UTILITY BONDS	YES	2
129	13026	YIELD ON HIGH GRADE INDUSTRIAL BONDS, AAA RATING	YES	2
130	13030	WEIGHTED AVERAGE OF OPEN MARKET RATES, NEW YORK CITY	YES	2
131	13031	BANK RATES ON CUSTOMER LOANS, LEADING CITIES	YES	2
132	13032	TOTAL RATES CHARGED CUSTOMERS AND OPEN MARKET RATES, COMBINED	YES	2
133	13033	YIELD ON LONG-TERM UNITED STATES BONDS	YES	2
134	13035	YIELDS ON CORPORATE BONDS, HIGHEST RATING	YES	2
135	13036	YIELDS ON CORPORATE BONDS, LOWEST RATING	YES	2
136	13048	DIVIDEND YIELD OF PREFERRED STOCK ON THE NEW YORK STOCK EXCHANGE	YES	2
137	14062	TOTAL GOLD RESERVES OF FEDERAL RESERVE BANKS	YES	2
138	14063	CASH RESERVES OF FEDERAL RESERVE BANKS	YES	2
139	14064	RESERVES HELD AT FEDERAL RESERVE BANKS, ALL MEMBER BANKS	YES	2
140	14065	NOTES IN CIRCULATION, FEDERAL RESERVE BANKS	YES	2
141	14066	TOTAL BILLS AND SECURITIES HELD BY FEDERAL RESERVE BANKS	YES	2
142	14067	BILLS DISCOUNTED, FEDERAL RESERVE BANKS	YES	2
143	14069	GOVERNMENT SECURITIES HELD, FEDERAL RESERVE BANKS	YES	2
144	14070	TOTAL DEPOSITS, FEDERAL RESERVE BANKS	YES	2
145	14072	RATIO OF RESERVES TO NOTE AND DEPOSIT LIABILITIES, FEDERAL RESERVE BANKS	YES	2
146	14076	MONETARY GOLD STOCK	YES	2
147	14078	NET DEMAND DEPOSITS, REPORTING MEMBER BANKS, FEDERAL RESERVE SYSTEM	YES	2
148	14079	TIME DEPOSITS, REPORTING MEMBER BANKS, FEDERAL RESERVE SYSTEM	YES	2
149	14080	CURRENCY HELD BY THE TREASURY	YES	2
150	14086	PERCENTAGE OF RESERVES HELD TO RESERVES REQUIRED, ALL MEMBER BANKS, FRB SYSTEM	YES	2
151	14121	NEW YORK CITY	YES	2
152	14126	VAULT CASH, ALL BANKS EXCEPT FEDERAL RESERVE BANKS	YES	2
153	14127	INVESTMENTS OTHER THAN UNITED STATES GOVERNMENT SECURITIES, REPORTING FEDERAL RESERVE MEMBER BANKS IN 101 LEADING CITIES	YES	2
154	14135	TOTAL CURRENCY OUTSIDE THE TREASURY AND FEDERAL RESERVE BANKS, END OF MONTH	YES	2
155	14137	GOLD HELD IN THE TREASURY AND FEDERAL RESERVE BANKS, END OF	YES	2
156	14144	MONEY STOCK, COMMERICAL BANKS PLUS CURRENCY HELD BY PUBLIC	YES	2
157	14145	TOTAL DEPOSITS, ALL COMMERCIAL BANKS	YES	2
158	14172	ADJUSTED DEMAND DEPOSITS, ALL COMMERCIAL BANKS	YES	2
159	14173	DEPOSITS IN MUTUAL SAVINGS BANKS AND POSTAL SAVINGS SYSTEM, END OF MONTH	YES	2
160	14174	ADJ. DEMAND DEPOSITS, ALL COMMERCIAL BANKS,CURRENCY HELD BY PUBLIC	YES	2
161	14175	ADJ. DEMAND DEPOSITS, ALL BANKS,TOTAL TIME DEPOSITS, CURRENCY HELD BY PUBLIC	YES	2
162	14178	RATIO OF CURRENCY HELD BY THE PUBLIC TO ADJUSTED DEMAND DEPOSITS, TIME DEPOSITS, ALL COMMERCIAL BANKS, PLUS CURRENCY HELD BY THE PUBLIC	YES	2
163	14190	PERCENT CHANGE IN TOTAL MONEY SUPPLY, MONTH-TO-MONTH CHANGE	YES	2
164	14195	MONEY STOCK, MONTH-TO-MONTH CHANGE	YES	2

Appendix B.3 Tables

Table 3.4: **Imposed sign restriction to identify a contractionary monetary policy shock.**

Indicator series	**Sign Restriction**	**Horizon**
Consumer price index	-	6
Index of general price level	-	6
Whole sale price of metal and metal products	-	6
M1	-	6
Commercial paper rates	+	6
Discount rate (New York)	+	6

Imposed Sign restriction to identify a contractionary monetary policy shock.

Table 3.5: **Macro Variables and share of variance explained by estimated factors**

The table delivers the forecast error variance decomposition of a monetary poicy shock for the 3 models considered. The respective 3 blocks report the results for the discount rate model, commercial paper rates model and the M1 model. The variables considered are the same as for the impulse response analysis, namely DR is the Discount Rate, CPR is the commercial Paper rate, Y is the growth in FRB index for production in manufacturing, π is CPI inflation, S&P is the Standard and Poors 500 index, I stands for the index of orders in Machinery and Tools. The firs value denotes the median based on the posterior draws, and the following two values are the uncertainty covering the 68% highest posterior density.

Commercial Paper Rate Model								
Horizon	0	1	2	3	6	12	24	48
CommPR	5	6	7	7	6	6	5	5
FRB Industrial Production	5	6	8	8	8	9	9	10
CPI inflation	65	48	46	44	43	43	43	42
S&P 500	9	11	13	14	15	15	16	16
Wages	12	14	15	15	15	15	16	16
Orders of Machinery Tools	9	13	14	18	18	20	20	20
Discount Rate Model								
Horizon	0	1	2	3	6	12	24	48
Discount Rate	4	5	5	5	5	6	6	6
FRB Industrial Production	4	4	5	6	6	7	7	9
CPI inflation	94	56	53	52	48	46	45	43
S&P 500	5	6	6	7	8	9	9	10
Wages	10	11	11	11	11	12	12	13
Orders of Machinery Tools	6	7	7	9	10	10	11	11
M0 Model								
Horizon	0	1	2	3	6	12	24	48
M0	4	4	4	4	3	3	3	3
FRB Industrial Production	15	18	19	18	18	20	21	20
CPI inflation	54	38	38	37	37	37	36	35
S&P 500	16	18	20	22	22	22	22	21
Wages	10	10	12	12	12	13	12	12
Orders of Machinery Tools	18	17	20	23	23	23	23	22
M1 Model								
Horizon	0	1	2	3	6	12	24	48
M1	17	19	19	18	17	18	17	16
FRB Industrial Production	16	17	17	17	17	17	16	16
CPI inflation	89	60	55	54	50	48	47	42
S&P 500	19	20	21	22	22	22	23	23
Wages	16	17	18	18	17	17	16	16
Orders of Machinery Tools	22	23	26	27	26	26	26	27

Table 3.6: **RMSFE at horizon 1 for different models considered.**

Horizon: 1	FRB Ind. Prod.	CPI Inflation	Orders	Policy variable
		Commercial Paper Rate Model		
Before Great Crash	5.85,[12.01,2.36]	1.27,[2.66,0.69]	26.35,[41.49,15.50]	133.09,[145.70,129.35]
Before 1^{st} Banking Crisis	8.40,[16.42,2.85]	2.04,[3.80,0.90]	27.24,[38.82,17.58]	133.78,[148.89,129.46]
Before German Banking Crisis	4.59,[9.26,2.03]	1.07,[2.01,0.65]	20.62,[33.04,13.66]	130.47,[135.20,129.12]
Before Devaluation of £	3.23,[6.73,1.78]	0.91,[1.53,0.63]	20.05,[31.91,13.58]	130.23,[134.53,129.13]
Before "Banking Holiday"	2.67,[4.77,1.74]	0.75,[1.15,0.61]	16.33,[21.89,12.88]	131.01,[137.08,129.20]
		Discount Rate Model		
Before Great Crash	7.03,[9.79,4.83]	1.11,[1.58,0.93]	26.21,[32.67,21.52]	216.26,[218.28,215.25]
Before 1^{st} Banking Crisis	4.81,[8.17,3.64]	1.01,[1.47,0.91]	22.33,[26.70,20.50]	222.20,[227.63,218.69]
Before German Banking Crisis	5.31,[7.12,4.05]	0.96,[1.14,0.91]	22.57,[26.96,20.92]	227.06,[231.73,223.60]
Before Devaluation of £	6.05,[8.43,4.32]	0.98,[1.25,0.91]	24.44,[30.78,21.89]	229.63,[235.72,224.73]
Before "Banking Holiday"	5.32,[6.70,4.22]	0.96,[1.10,0.91]	22.04,[24.33,20.52]	225.14,[229.18,221.86]
		M0 Model		
Before Great Crash	2.88,[3.50,2.69]	1.03,[1.16,0.99]	22.32,[23.73,21.51]	2.36,[2.45,2.34]
Before 1^{st} Banking Crisis	2.82,[3.35,2.69]	1.02,[1.11,0.99]	22.89,[24.05,22.09]	2.36,[2.45,2.34]
Before German Banking Crisis	2.83,[3.22,2.68]	1.02,[1.11,0.99]	24.32,[26.97,22.28]	2.52,[2.69,2.40]
Before Devaluation of £	2.88,[3.42,2.69]	1.03,[1.14,0.99]	24.95,[27.34,23.23]	2.47,[2.65,2.37]
Before "Banking Holiday"	3.20,[4.04,2.74]	1.02,[1.14,0.99]	22.99,[25.03,21.98]	2.49,[2.85,2.35]
		M1 Model		
Before Great Crash	2.48,[2.87,2.38]	0.81,[0.96,0.76]	24.33,[25.32,23.76]	3.73,[3.76,3.72]
Before 1^{st} Banking Crisis	2.55,[3.08,2.39]	0.83,[1.03,0.77]	25.26,[26.32,24.38]	3.76,[3.88,3.72]
Before German Banking Crisis	2.55,[3.15,2.39]	0.82,[0.99,0.77]	25.91,[27.98,24.72]	3.75,[3.87,3.72]
Before Devaluation of £	2.49,[2.83,2.38]	0.82,[0.95,0.77]	24.46,[26.35,23.21]	3.75,[3.83,3.72]
Before "Banking Holiday"	2.51,[3.00,2.38]	0.92,[1.12,0.78]	23.68,[25.29,22.94]	3.75,[3.84,3.72]
		No Policy Instrument Model		
Before Great Crash	5.72,[6.18,5.63]	3.02,[3.71,2.70]	2.94,[4.96,2.17]	
Before 1^{st} Banking Crisis	4.96,[5.84,4.52]	7.50,[7.89,7.40]	3.73,[5.02,2.91]	
Before German Banking Crisis	4.97,[5.28,4.89]	7.45,[7.90,7.25]	2.90,[4.06,2.37]	
Before Devaluation of £	4.11,[4.67,3.95]	10.90,[12.15,9.67]	3.23,[4.69,2.53]	
Before "Banking Holiday"	5.04,[5.39,4.94]	5.58,[6.23,5.23]	2.77,[4.11,2.05]	

The table reports the median root mean squared forecast errors (RMSFE) and its 68% highest posterior density in brackets for the respective models under consideration.

Table 3.7: **RMSFE at horizon 3 for different models considered.**

Horizon: 3	FRB Ind. Prod.	CPI Inflation	Orders	Policy variable
		Commercial Paper Rate Model		
Before Great Crash	10.25,[18.37,5.08]	2.27,[4.19,1.25]	26.62,[41.01,16.22]	175.41,[277.21,127.55]
Before 1^{st} Banking Crisis	13.40,[22.12,6.76]	2.86,[4.69,1.57]	29.32,[41.70,18.86]	173.18,[280.22,128.50]
Before German Banking Crisis	7.02,[13.53,3.43]	1.60,[2.87,0.92]	21.16,[33.21,13.88]	137.95,[187.04,122.88]
Before Devaluation of £	5.33,[10.14,2.84]	1.37,[2.31,0.82]	19.91,[31.60,13.43]	131.54,[172.45,122.19]
Before "Banking Holiday"	4.58,[7.70,2.60]	1.28,[1.92,0.85]	17.64,[23.50,13.65]	162.30,[230.61,127.54]
		Discount Rate Model		
Before Great Crash	8.09,[11.60,5.43]	1.46,[2.26,1.04]	25.03,[31.20,20.66]	207.81,[218.92,204.04]
Before 1^{st} Banking Crisis	5.54,[9.91,4.06]	1.20,[1.99,0.95]	21.51,[26.03,19.64]	217.85,[239.00,207.57]
Before German Banking Crisis	5.43,[7.39,4.26]	1.04,[1.36,0.91]	21.83,[26.03,20.05]	236.14,[254.55,222.72]
Before Devaluation of £	6.26,[9.06,4.73]	1.14,[1.62,0.93]	23.67,[29.80,21.08]	244.87,[269.52,225.58]
Before "Banking Holiday"	5.35,[6.95,4.33]	1.02,[1.28,0.90]	21.27,[23.85,19.61]	231.85,[249.14,218.80]
		M0 Model		
Before Great Crash	3.17,[4.05,2.72]	1.07,[1.26,0.97]	22.91,[23.97,22.14]	2.26,[2.38,2.21]
Before 1^{st} Banking Crisis	3.11,[4.16,2.67]	1.04,[1.24,0.96]	22.23,[23.25,21.58]	2.26,[2.39,2.20]
Before German Banking Crisis	3.00,[3.53,2.70]	1.02,[1.17,0.95]	23.28,[25.59,21.71]	2.40,[2.59,2.28]
Before Devaluation of £	3.15,[3.80,2.78]	1.05,[1.23,0.96]	23.88,[25.90,22.55]	2.41,[2.61,2.28]
Before "Banking Holiday"	3.55,[4.67,2.86]	1.08,[1.29,0.97]	22.47,[24.09,21.60]	2.63,[3.12,2.33]
		M1 Model		
Before Great Crash	3.34,[4.45,2.61]	0.92,[1.15,0.79]	26.02,[27.90,24.85]	3.50,[3.56,3.47]
Before 1^{st} Banking Crisis	3.61,[4.94,2.80]	0.96,[1.18,0.81]	28.84,[30.39,27.50]	3.61,[3.79,3.51]
Before German Banking Crisis	3.27,[4.41,2.58]	0.90,[1.13,0.77]	29.59,[31.86,27.61]	3.62,[3.82,3.52]
Before Devaluation of £	3.09,[3.77,2.59]	0.85,[0.99,0.76]	28.01,[30.11,26.34]	3.55,[3.66,3.49]
Before "Banking Holiday"	3.43,[4.76,2.67]	0.94,[1.16,0.80]	26.87,[29.12,25.49]	3.56,[3.67,3.50]
		No Policy Instrument Model		
Before Great Crash	5.67,[6.27,5.37]	3.19,[4.03,2.76]	3.04,[4.98,2.21]	
Before 1^{st} Banking Crisis	5.13,[6.16,4.53]	7.24,[7.81,7.00]	3.65,[4.88,2.80]	
Before German Banking Crisis	5.52,[6.49,4.90]	11.41,[12.48,10.40]	3.12,[4.24,2.58]	
Before Devaluation of £	4.33,[5.07,3.91]	10.32,[11.48,9.19]	3.57,[4.94,2.87]	
Before "Banking Holiday"	5.11,[5.61,4.84]	6.33,[7.20,5.67]	2.95,[4.18,2.21]	

The table reports the median root mean squared forecast errors (RMSFE) and its 68% highest posterior density in brackets for the respective models under consideration.

Table 3.8: **RMSFE at horizon 6 for different models considered.**

Horizon: 6	FRB Ind. Prod.	CPI Inflation	Orders	Policy variable
		Commercial Paper Rate Model		
Before Great Crash	12.30,[20.24,7.05]	3.43,[5.87,1.94]	28.70,[44.83,17.80]	289.42,[542.88,150.07]
Before 1^{st} Banking Crisis	14.84,[22.48,8.92]	3.31,[4.96,2.12]	29.57,[42.01,19.50]	285.84,[545.34,152.04]
Before German Banking Crisis	8.15,[13.96,4.58]	2.03,[3.40,1.21]	21.27,[32.70,13.99]	191.07,[350.66,124.44]
Before Devaluation of £	6.72,[11.28,3.80]	1.68,[2.83,1.02]	20.31,[31.42,13.67]	166.75,[301.88,119.88]
Before "Banking Holiday"	5.64,[8.82,3.55]	2.22,[3.24,1.43]	19.54,[27.56,14.41]	257.96,[412.17,147.00]
		Discount Rate Model		
Before Great Crash	8.43,[11.62,6.04]	1.79,[2.79,1.19]	24.53,[30.89,20.29]	227.52,[280.63,196.91]
Before 1^{st} Banking Crisis	6.24,[11.07,4.50]	1.35,[2.45,0.99]	21.78,[26.41,19.68]	209.59,[262.18,194.64]
Before German Banking Crisis	6.35,[8.21,5.04]	1.14,[1.52,0.95]	22.82,[26.73,20.43]	229.69,[262.91,209.85]
Before Devaluation of £	7.41,[10.01,5.72]	1.30,[1.93,1.00]	24.83,[30.57,21.74]	242.30,[287.81,213.92]
Before "Banking Holiday"	6.06,[7.54,4.97]	1.11,[1.41,0.93]	22.20,[25.34,20.06]	229.05,[257.82,207.66]
		M0 Model		
Before Great Crash	3.90,[4.99,3.16]	1.15,[1.46,0.99]	24.44,[25.63,23.73]	2.26,[2.44,2.14]
Before 1^{st} Banking Crisis	3.97,[5.39,3.08]	1.19,[1.52,1.00]	25.43,[26.70,24.37]	2.22,[2.42,2.11]
Before German Banking Crisis	3.32,[4.03,2.86]	1.08,[1.31,0.95]	26.37,[28.55,24.71]	2.26,[2.45,2.14]
Before Devaluation of £	3.77,[4.65,3.09]	1.15,[1.38,0.98]	27.20,[29.19,25.62]	2.26,[2.46,2.13]
Before "Banking Holiday"	4.42,[5.60,3.53]	1.21,[1.45,1.03]	25.37,[27.19,24.24]	2.61,[3.10,2.27]
		M1 Model		
Before Great Crash	3.59,[4.69,2.81]	1.10,[1.44,0.87]	25.38,[26.62,24.56]	3.25,[3.33,3.20]
Before 1^{st} Banking Crisis	4.08,[5.47,3.14]	1.29,[1.64,1.03]	26.85,[28.30,25.66]	3.49,[3.73,3.34]
Before German Banking Crisis	3.53,[4.61,2.83]	1.04,[1.30,0.86]	27.55,[29.62,25.82]	3.40,[3.61,3.28]
Before Devaluation of £	3.34,[4.11,2.74]	0.95,[1.15,0.81]	26.13,[28.00,24.69]	3.33,[3.46,3.25]
Before "Banking Holiday"	3.77,[5.04,2.93]	1.06,[1.31,0.86]	25.23,[27.11,24.26]	3.35,[3.49,3.26]
		No Policy Instrument Model		
Before Great Crash	5.54,[6.26,5.12]	3.81,[4.97,3.15]	3.17,[5.11,2.30]	
Before 1^{st} Banking Crisis	5.36,[6.55,4.58]	7.01,[7.86,6.63]	3.62,[4.83,2.80]	
Before German Banking Crisis	5.34,[6.28,4.72]	10.59,[11.64,9.65]	3.47,[4.65,2.83]	
Before Devaluation of £	4.37,[5.12,3.85]	9.62,[10.70,8.58]	3.43,[4.78,2.75]	
Before "Banking Holiday"	5.24,[5.88,4.79]	7.15,[8.03,6.38]	3.33,[4.44,2.68]	

The table reports the median root mean squared forecast errors (RMSFE) and its 68% highest posterior density in brackets for the respective models under consideration.

Table 3.9: **RMSFE at horizon 12 for different models considered.**

Horizon: 12	FRB Ind. Prod.	CPI Inflation	Orders	Policy variable
		Commercial Paper Rate Model		
Before Great Crash	13.56,[20.31,9.14]	4.07,[6.61,2.51]	33.09,[52.76,20.59]	457.83,[903.31,217.92]
Before 1^{st} Banking Crisis	15.49,[22.26,10.58]	3.52,[5.04,2.46]	31.20,[45.39,21.72]	463.94,[891.51,217.75]
Before German Banking Crisis	8.94,[13.76,5.68]	2.38,[3.73,1.53]	22.91,[34.13,15.93]	322.01,[632.84,155.79]
Before Devaluation of £	7.35,[11.43,4.60]	2.03,[3.15,1.31]	21.52,[31.66,14.99]	268.69,[558.88,141.01]
Before "Banking Holiday"	6.95,[10.22,4.66]	2.58,[3.64,1.70]	25.61,[37.56,17.56]	337.30,[582.45,171.61]
		Discount Rate Model		
Before Great Crash	8.55,[11.55,6.37]	1.99,[2.91,1.40]	26.35,[32.83,22.68]	269.33,[390.45,196.51]
Before 1^{st} Banking Crisis	7.14,[11.58,5.54]	1.53,[2.80,1.10]	25.29,[30.96,22.75]	212.80,[370.87,181.15]
Before German Banking Crisis	6.79,[8.29,5.68]	1.41,[1.85,1.12]	26.29,[30.48,23.67]	218.65,[263.32,196.30]
Before Devaluation of £	7.49,[9.74,6.07]	1.60,[2.31,1.21]	27.59,[33.20,24.49]	240.79,[323.19,206.28]
Before "Banking Holiday"	7.00,[8.29,5.98]	1.40,[1.73,1.15]	26.75,[30.33,24.05]	212.68,[245.60,191.29]
		M0 Model		
Before Great Crash	3.96,[5.04,3.24]	1.26,[1.64,1.03]	23.71,[24.88,22.95]	2.10,[2.30,1.96]
Before 1^{st} Banking Crisis	4.43,[5.75,3.60]	1.59,[2.08,1.24]	24.33,[25.84,23.32]	2.09,[2.30,1.95]
Before German Banking Crisis	3.37,[4.05,2.87]	1.23,[1.54,1.00]	24.66,[26.54,23.29]	2.15,[2.39,1.99]
Before Devaluation of £	3.88,[4.84,3.21]	1.37,[1.65,1.14]	25.42,[27.24,24.02]	2.20,[2.46,2.02]
Before "Banking Holiday"	4.66,[5.78,3.77]	1.36,[1.66,1.12]	24.35,[26.11,23.19]	2.68,[3.17,2.31]
		M1 Model		
Before Great Crash	3.65,[4.69,2.92]	1.16,[1.50,0.92]	24.43,[25.91,23.61]	2.93,[3.06,2.85]
Before 1^{st} Banking Crisis	4.39,[5.66,3.54]	1.47,[1.84,1.20]	25.84,[27.49,24.66]	3.40,[3.71,3.19]
Before German Banking Crisis	3.70,[4.71,3.00]	1.16,[1.45,0.94]	26.13,[28.22,24.56]	3.18,[3.42,3.02]
Before Devaluation of £	3.54,[4.31,2.95]	1.05,[1.25,0.87]	25.09,[26.78,23.68]	3.08,[3.24,2.96]
Before "Banking Holiday"	4.15,[5.33,3.29]	1.15,[1.44,0.92]	24.27,[26.30,23.14]	3.14,[3.35,2.98]
		No Policy Instrument Model		
Before Great Crash	5.44,[6.24,4.91]	3.92,[5.16,3.22]	3.44,[5.41,2.56]	
Before 1^{st} Banking Crisis	5.84,[7.14,4.92]	9.12,[10.24,8.31]	4.02,[5.44,3.19]	
Before German Banking Crisis	5.01,[5.83,4.41]	9.36,[10.25,8.54]	3.35,[4.72,2.64]	
Before Devaluation of £	4.52,[5.29,3.91]	8.57,[9.56,7.67]	3.53,[4.86,2.79]	
Before "Banking Holiday"	5.06,[5.68,4.60]	6.55,[7.36,5.85]	3.54,[4.58,2.85]	

The table reports the median root mean squared forecast errors (RMSFE) and its 68% highest posterior density in brackets for the respective models under consideration.

Appendix B.4 Figures

Figure 3.3: **IRF of commercial paper rates model identified witrh sign restriction**

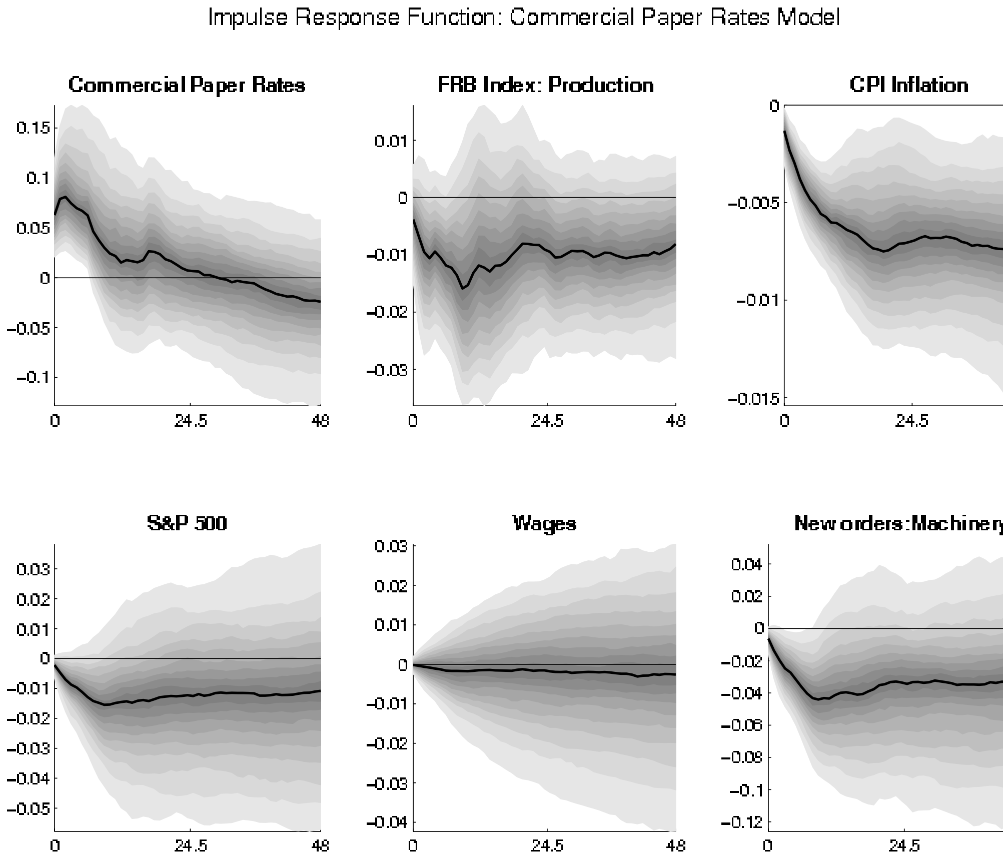


Figure 3.4: **IRF of discount rates model identified witrh sign restriction**

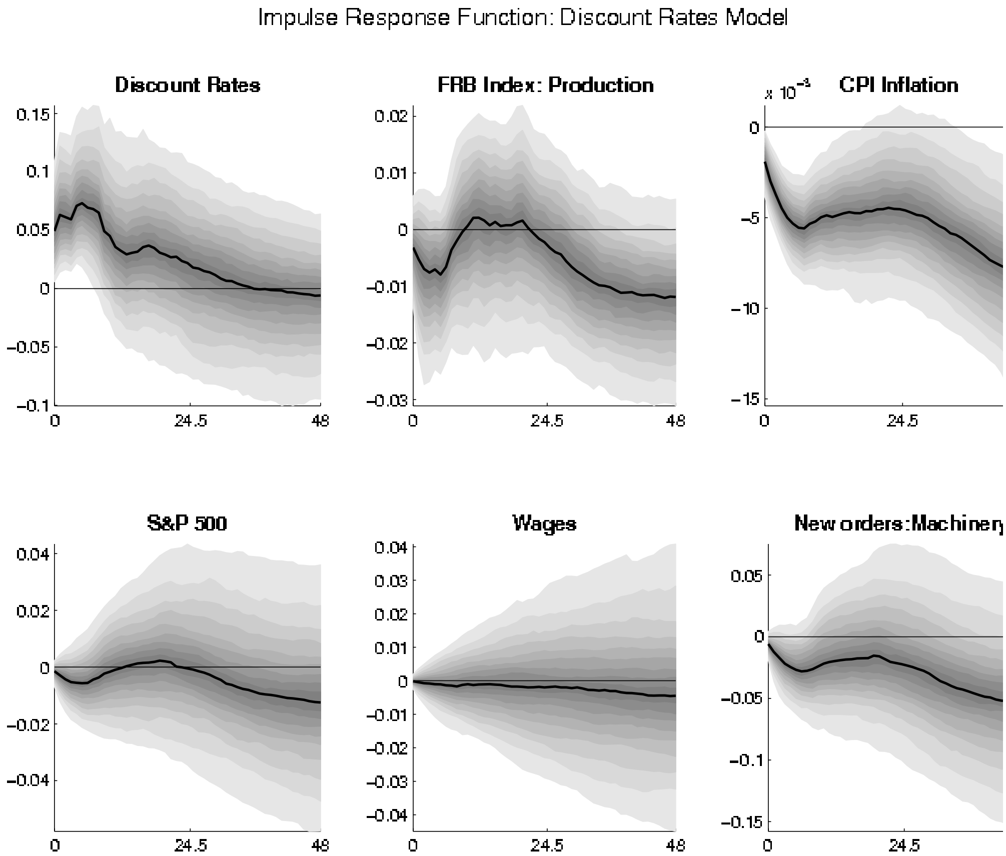


Figure 3.5: **IRF of M0 model identified witrh sign restriction**

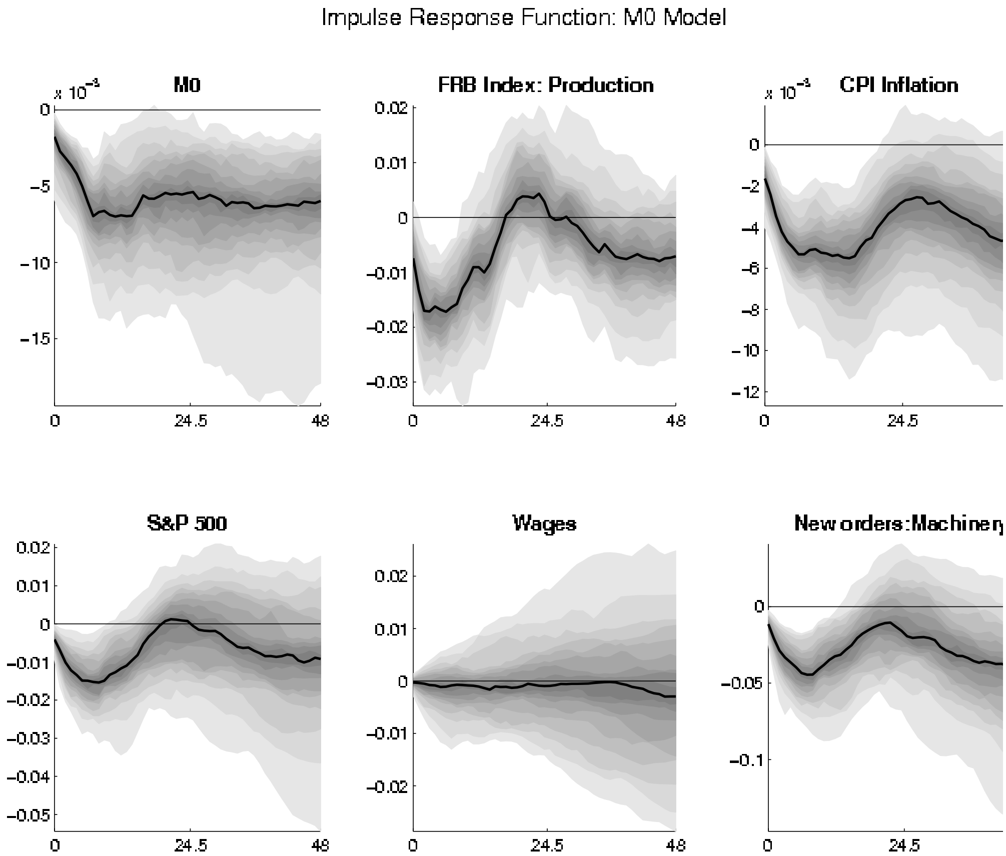


Figure 3.6: **IRF of M1 model identified witrh sign restriction**

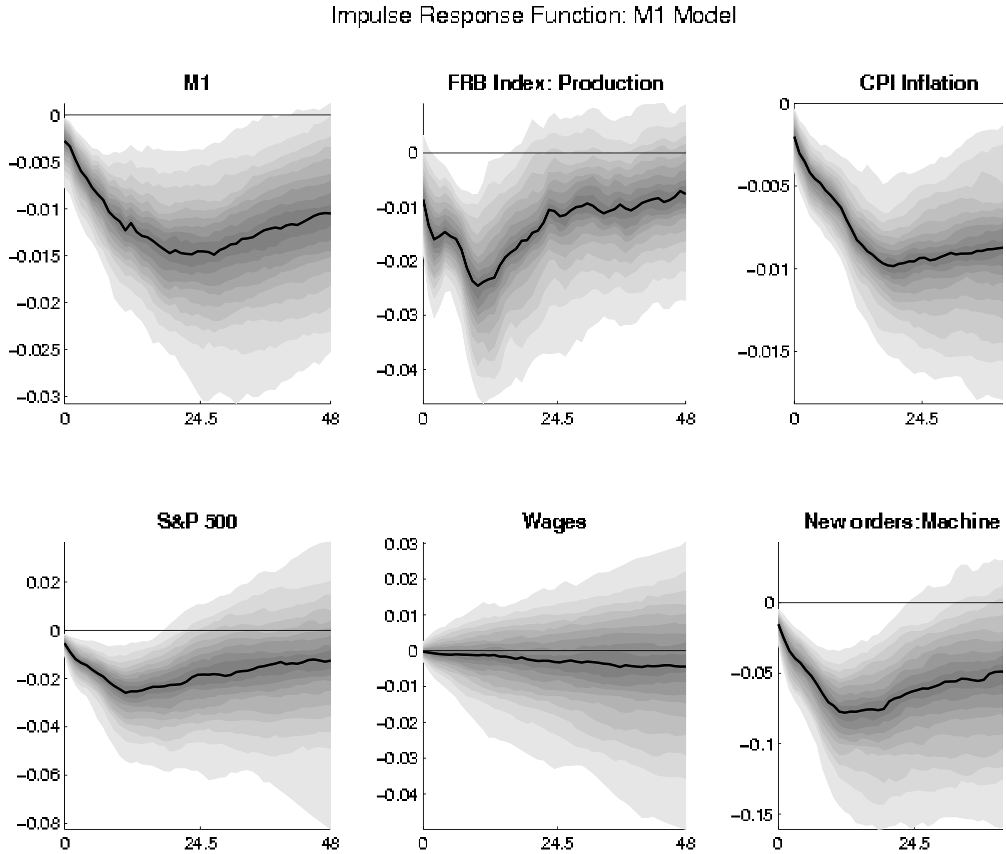


Figure 3.7: **IRF of M2 model identified witrh sign restriction**

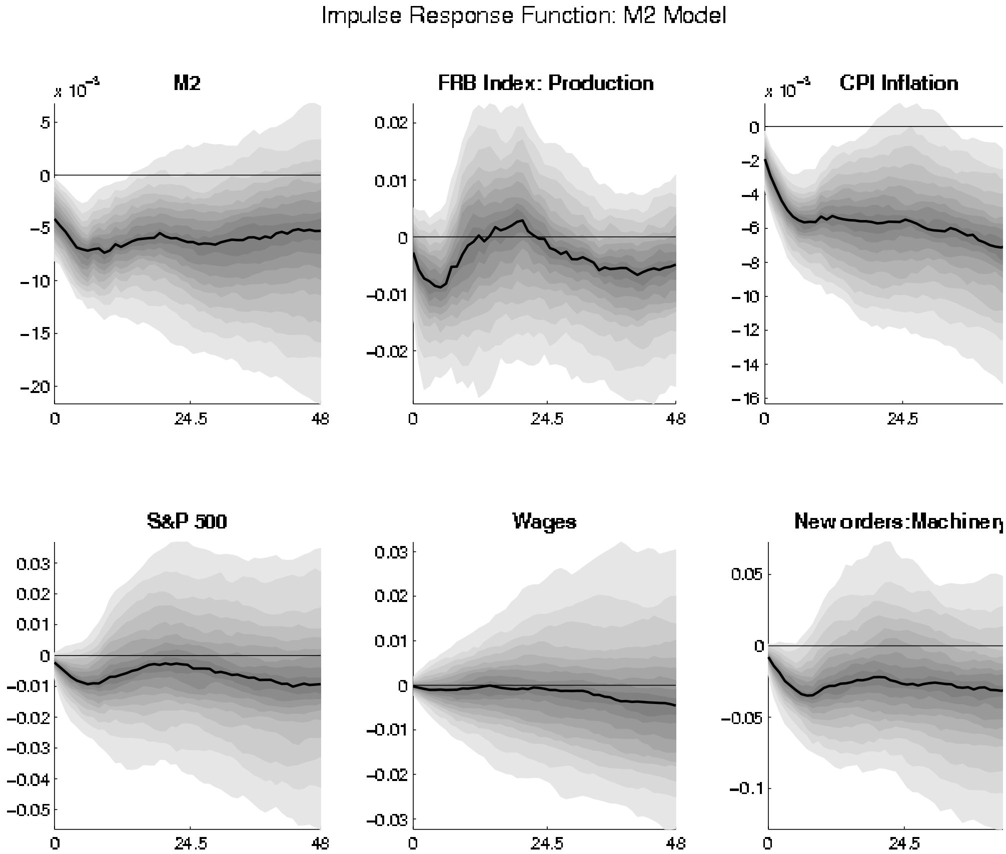


Figure 3.8: **Forecasting WITHOUT policy instruments**

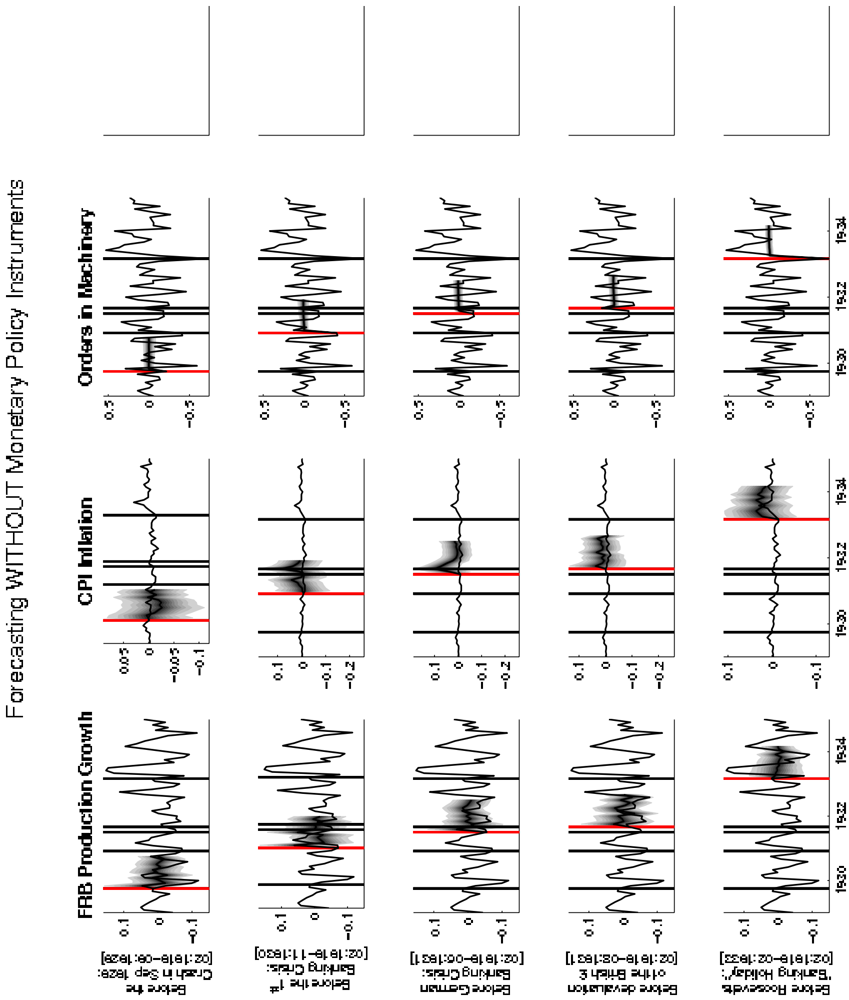


Figure 3.9: **Forecasting: Commercial Paper Rate Model**

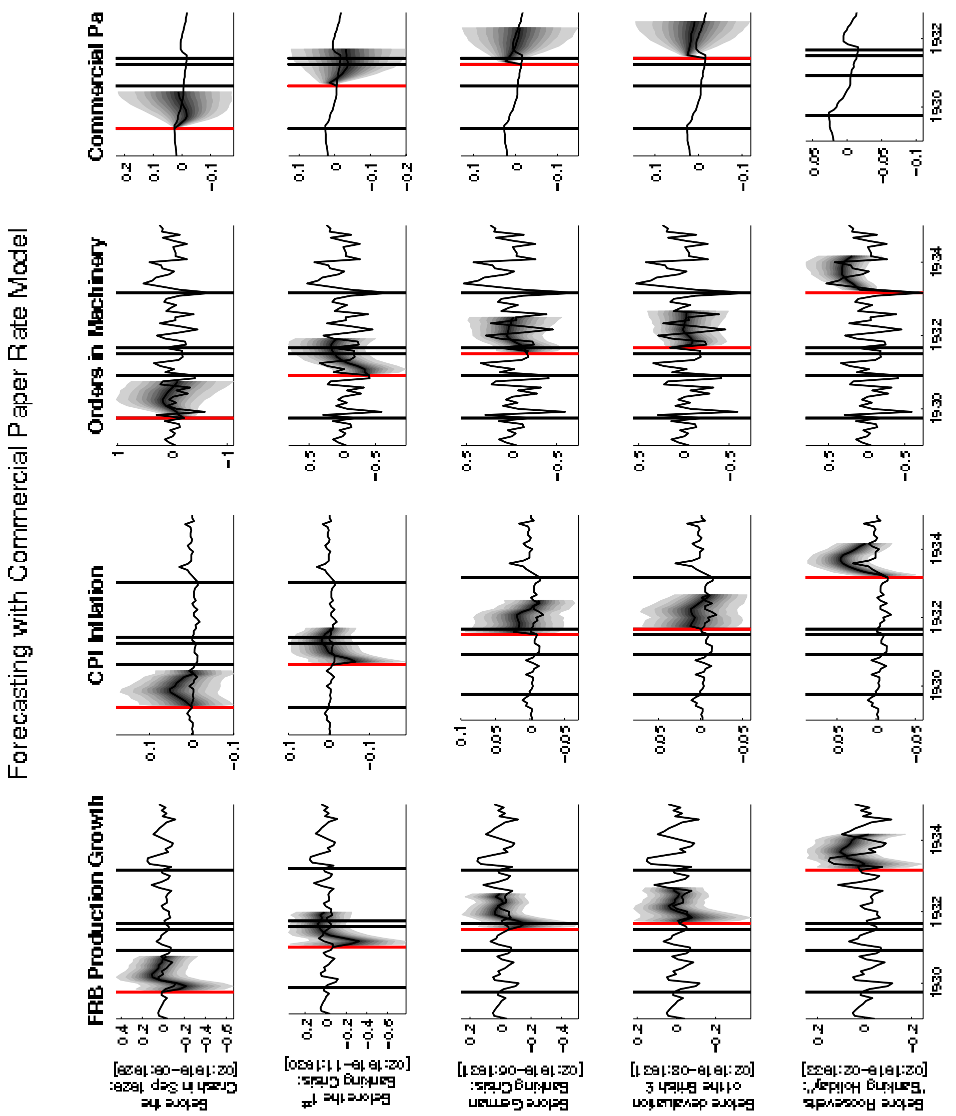
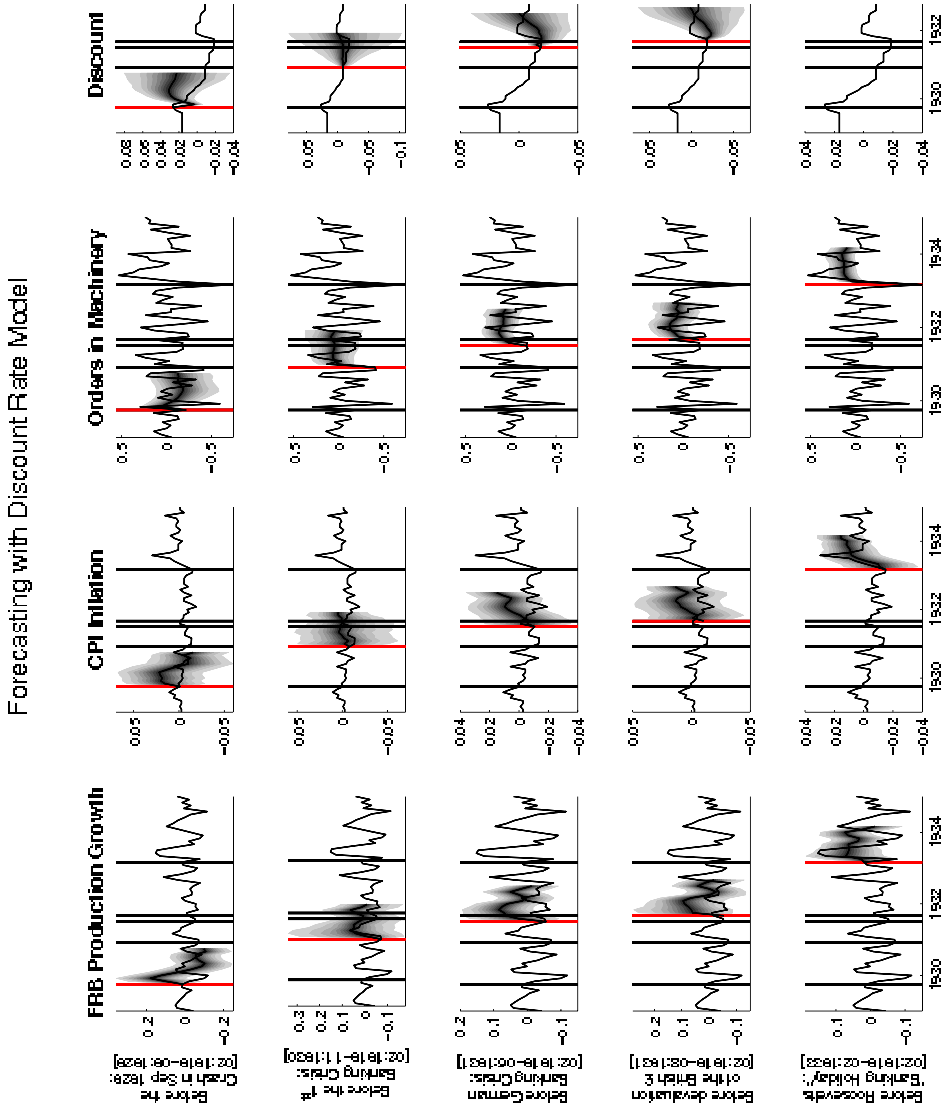


Figure 3.10: **Forecasting: Discount Rate Model**



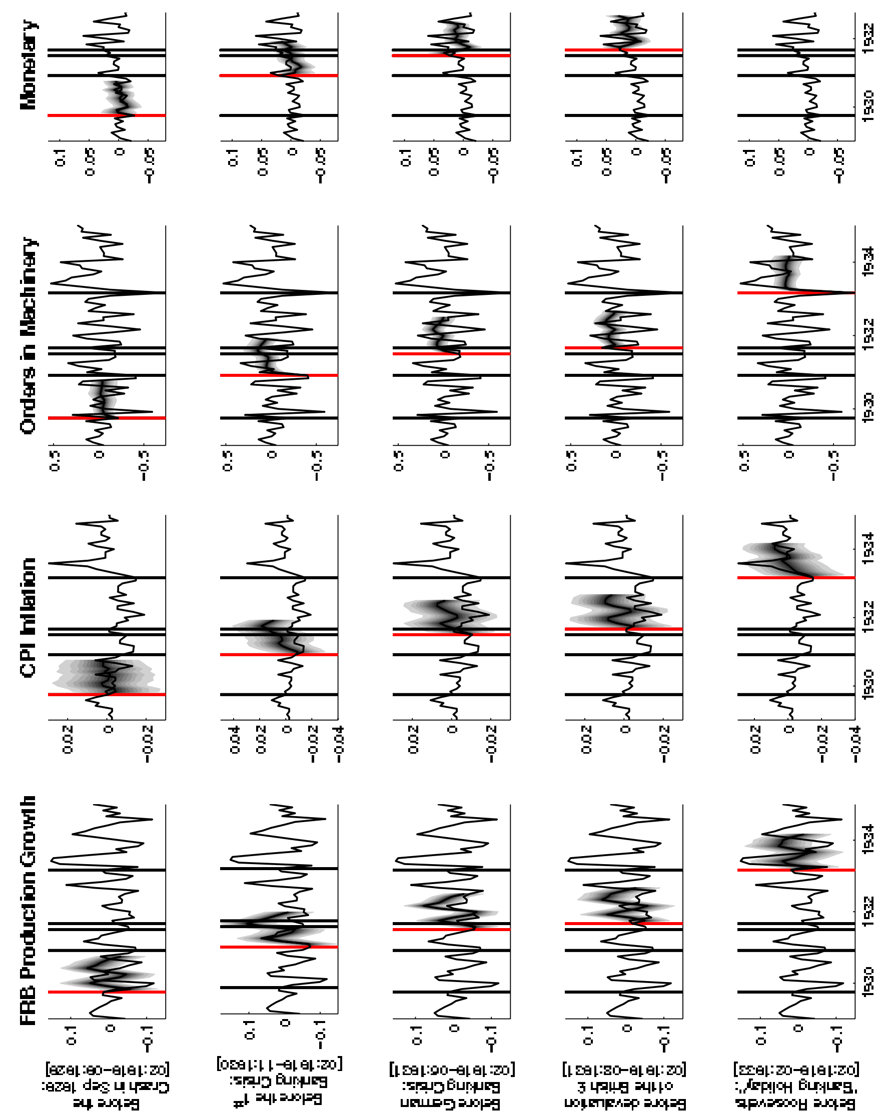


Figure 3.11: **Forecasting: M0 Model**

Forecasting with Monetary Base Model

Figure 3.12: **Forecasting: M1 Model**

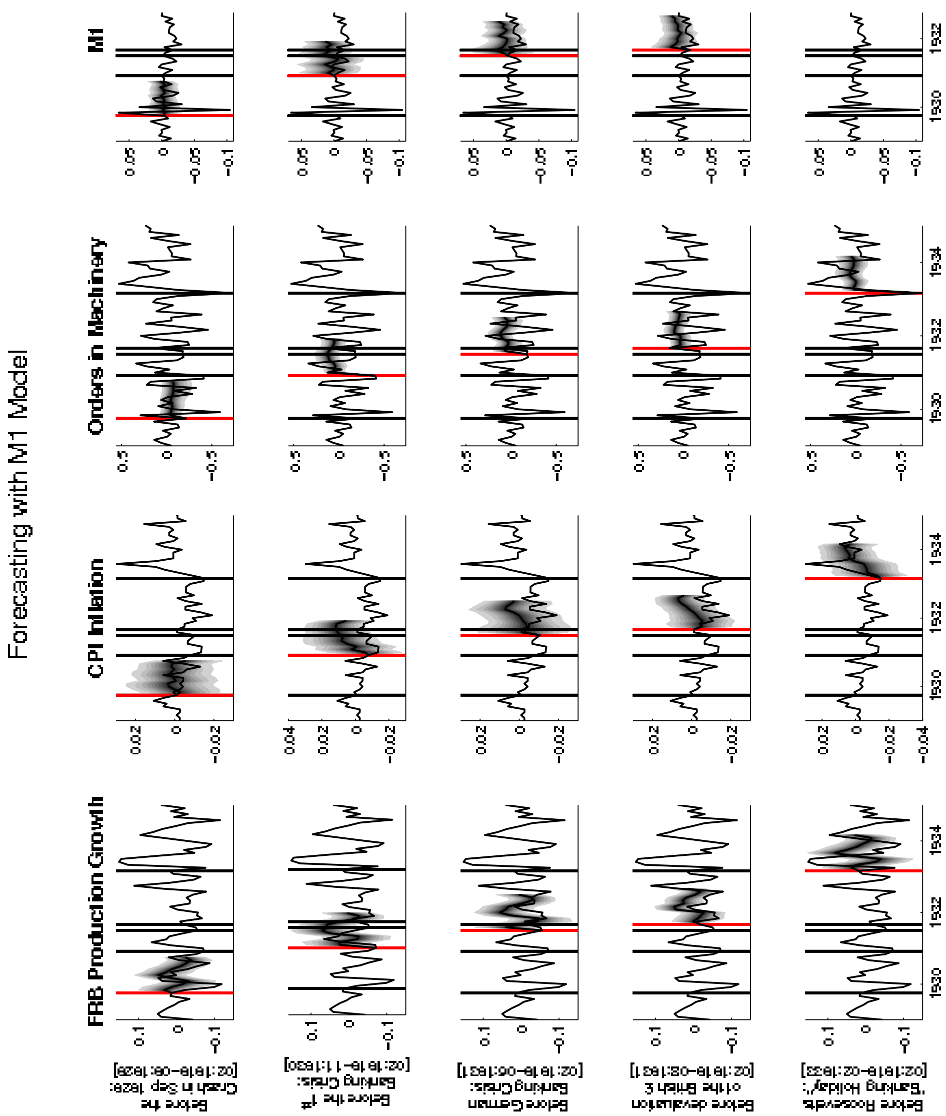


Figure 3.13: **Forecasting: M2 Model**

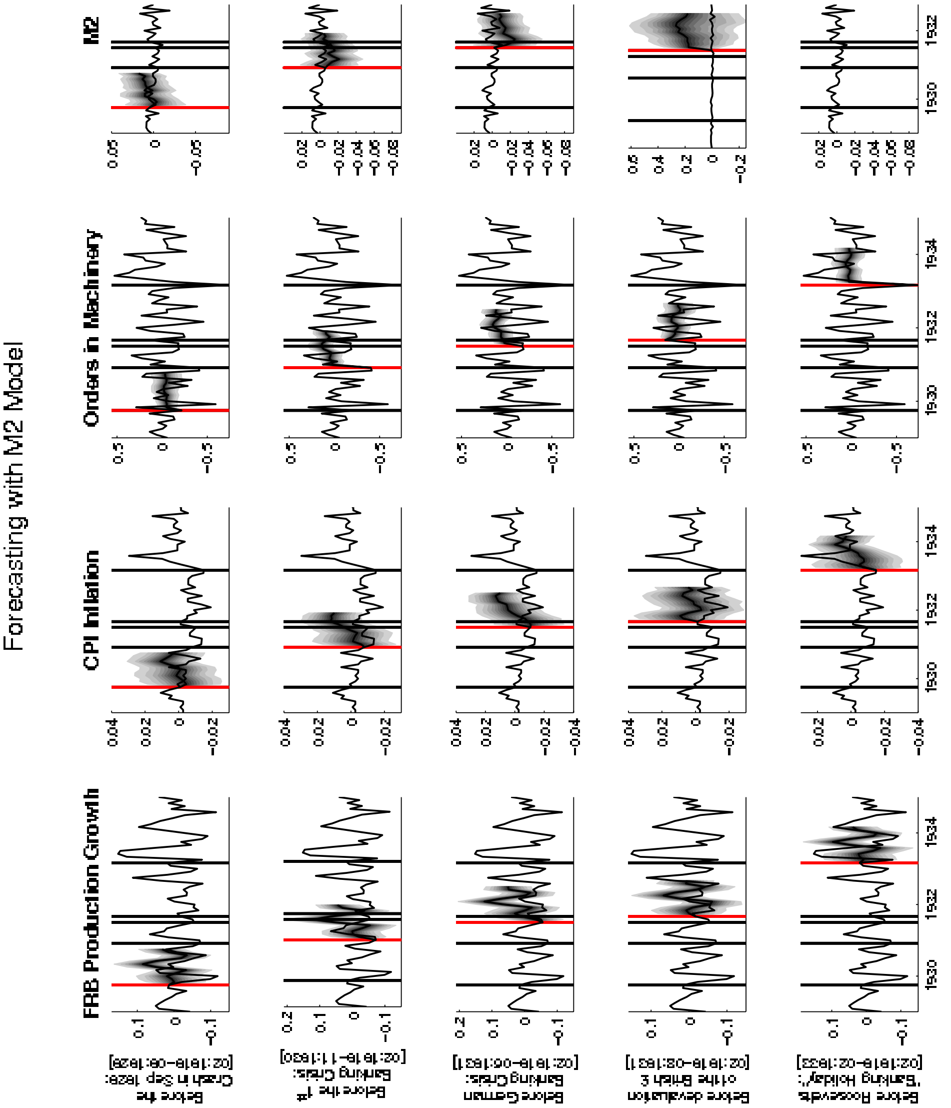
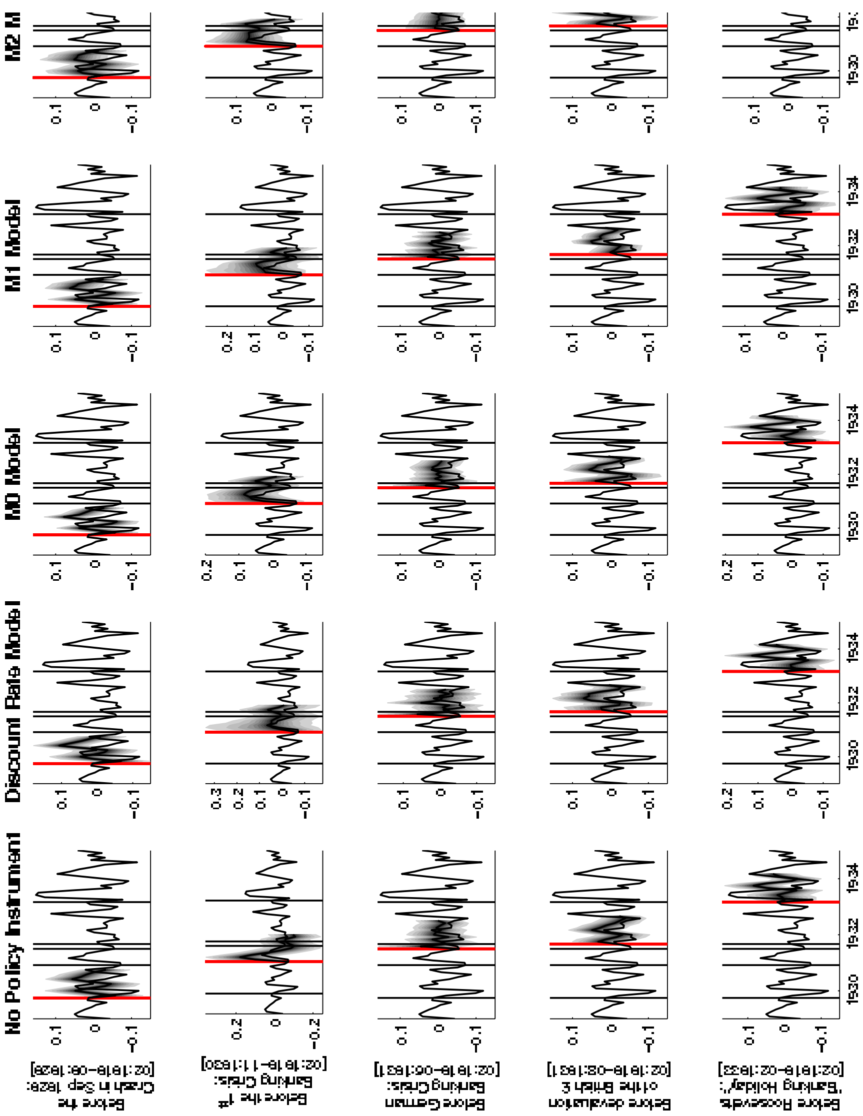


Figure 3.14: **Forecasting: OUTPUT INDICATORS**

Forecasting FRB Production Index in Growth Rates



4 Macroeconomic dynamics and shocks in the Euro area: Evidence from DSGE model-based identification in a Data-Rich Environment

This paper analyzes macroeconomic dynamics within the Euro area by first estimating the degree of macroeconomic comovement among constituent countries. Second I identify the effects of aggregate macroeconomic shocks across the member countries to detect heterogeneity in the transmission mechanism. To tackle the questions at hand I propose a novel approach to jointly estimate a factor-based DSGE model and a structural dynamic factor model that simultaneously captures the rich interrelations in a parsimonious way and explicitly involves economic theory in the estimation procedure. To identify shocks I employ both sign restrictions derived from the estimated DSGE model and the implied restrictions from the DSGE model rotation. I find a high degree of comovement across the member countries, homogeneity in the monetary transmission mechanism and heterogeneity in transmission of technology shocks. The suggested approach results in a factor generalization of the DSGE-VAR methodology of Del Negro and Schorfheide [12].

4.1 Introduction

What is the degree of comovement in the Euro area and is there heterogeneity in the transmission of aggregate shocks across the member countries of the currency union? There is little evidence on macroeconomic dynamics in the Euro area and even less about how aggregate shocks propagate across the member countries. Understanding the common dynamics and the propagation of the transmission mechanism of shocks across countries is important to understand for policy makers. The higher the degree of comovement the higher the exposure of its constituent countries to aggregate shocks however there might be strong differences in the exposure and in the transmission mechanism of shocks. Detecting and understanding the degree of homogeneity or heterogeneity in monetary transmission mechanism is particularly important and has important implications for the conduct of the single monetary policy and the different independent fiscal policies.

In this paper I propose to analyze the question at hand by jointly estimating a DSGE model as a prior generating tool to estimate and identify a Bayesian dynamic factor model (DFM). Different identification strategies are introduced within the proposed framework relying on DSGE models that also take the data-rich environment explicitly into account. Hence the joint estimation allows to efficiently matching theory to data in a flexible and parsimonious way. The theoretical model that I employ is a slightly modified version of the DSGE model estimated by Smets and Wouters for the Euro area. This choice is a natural one as it aims and succeeds to fit the data well though for the approach any other model of interest could be chosen. In addition this approach allows testing the validation of the models and the degree of misspecification can be determined.

I combine the strength of several recent advances in empirical macroeconomics in a unified and coherent framework to analyze the two questions at hand. Furthermore I extend and modify the respective approaches by proposing a factor generalization to parsimoniously exploit the informative content of large data sets available and important for unbiased inference. To be explicit my approach builds on four recent advances, namely (i) dynamic factor models (henceforth DFM) and Factor-Augmented VAR (henceforth FAVAR), (ii) the DSGE-VAR framework by Del Negro and Schorfheide [12] who use an estimated DSGE model to generate dummy observation for the estimation of a vector autoregression (VAR), (iii) the factor-based DSGE estimation approach of Boivin and Giannoni , (henceforth DSGE-F) and (iv) sign restrictions to identify macroeconomic shocks introduced by Uhlig [31], and Amir Ahmadi and Uhlig [2] for a factor generalization of the sign restriction identification scheme. The four approaches shall be explained here briefly.

The macroeconomic dynamics of the Euro area and its transmission mechanism of shocks have been analyzed in some studies. The closest is the paper by Boivin, Giannoni and Mojon who analyze the change in the monetary transmission mechanism of the Euro area since the introduction of the Euro. To carry out the exercise they employ a FAVAR to estimate impulse response function to a monetary policy shock and provides a structural interpretation of their findings by means of an open-economy DSGE model. My approach here is different in several dimensions. Most notably I propose different identification schemes as the recursive Cholesky identification Boivin, Giannoni and Mojon employ comes with several problems in particular present in large-dimensional models such as the FAVAR framework (see Amir Ahmadi and Uhlig [2] for a discussion and comparison with Sign restrictions). Furthermore I combine analysis with the joint estimation of a DSGE model with has a better fit. Regarding the methodology a paper related to the approach in this paper is Bäuerle who also combines a DFM with a DSGE model. In his paper he takes the New Keynesian Monetary DSGE model analyzed by Del Negro and Schorfheide [12] and combines it with a dynamic factor model on US data for interest rates, output and inflation.

Dynamic Factor Models. Since the early work of Sargent and Sims [23] and Geweke [19], dynamic factor analysis has gained increasing attention as a popular tool in empirical macroeconomics to represent the crucial dynamics of large data sets that can be decomposed in a common and an idiosyncratic component in a parsimonious way. The common factors describe the dynamic evolution of the data that is common across all indicators considered and the idiosyncratic component is variable specific. This approach holds the enticing promise to cope with and detect the rich interrelations of the plethora of data available to economic agents and policy makers prior to taking decisions. Alternative approaches relying on the standard VAR framework soon arrive at limitations due to the curse of dimensionality raised by Sims [25]. Disregarding data that is available to economic agents and policy makers leads to the well known *omitted variable bias* which can be successfully avoided by employing dynamic factor analysis[1]. Furthermore inference based on DFM turns out to be more robust compared to the VAR framework as regards the disturbing influence of measurement errors and the idiosyncratic components.

To facilitate estimation of the model I follow a Bayesian approach relying on posterior sampling based on Markov chain Monte Carlo simulation methods (henceforth MCMC). Here the potentially large dimensional parameter space and the unobserved factors in the state-space form can be split into different blocks from which we can sample directly or construct a proposal distribution (for a survey of these methods see Kim and Nelson [21]). As the factors are unobserved components, for the estimation procedure we have to sample the factor conditional on the blocks of parameters and the respective blocks of parameters conditional on the sampled latent factors. Additional complexity comes from the estimation of DSGE model which requires to sample the DSGE parameters from a nonstandard distribution hence we have to invoke a Metropolis-within-Gibbs sampling approach. The details can be found in the appendix.

Theory based identification. Key to the analysis of the transmission mechanism of aggregate shocks in a data-rich environment is the identification of shocks. Identifying the dynamic effects of an unanticipated change in one economic variable on other variables is of particular interest to macroeconomist and policy practioneers alike. There is a rich and well understood literature on identification of shocks within the VAR literature such as the long-run restriction , recursive identification , zero restrictions , sign restrictions, a formal approach building on explicit priors for impulse response function relying on probabilistic restriction and identified restrictions derived from estimated DSGE model rotation in the DSGE-VAR framework proposed by Del Negro and Schorfheide [12]. Thorough surveys can be found in Christiano, Eichenbaum and Evans. However studies focusing on the identification in large dimensional models like DFM and FAVAR are scarce. Amir Ahmadi and Uhlig [2] propose a factor generalization to identify monetary policy shocks within a FAVAR framework by exploiting the promising possibility to impose sign restriction on a larger though reasonably chosen according to *conventional wisdom*. They show that identifying monetary policy shocks in a FAVAR for the US economy outperforms the recursive Cholesky identification and the anomalies inherent which leads to unreasonable results in particular when controlling for changes in the monetary policy. In particular despite *weak* restriction it imposes the uncertainty associated is rather low.

Sign Restriction. The sign restriction approach is build on the idea to derive restriction from economic theory and directly impose it on the sign of the impulse response function of some variables while leaving others in particular those of specific interest unrestricted. Therefore sometimes this approach is referred to as agnostic identification. As one approach I propose to estimate a DSGE model that fits the data well such as the Smets and Wouters model for the Euro area and derive robust sign restrictions for the shocks of interest to be identified. These can be imposed on the impulse response functions of the factors if a structural version is specified or more flexibly on the single indicators of the DFM to identify the structural shocks. The advantage of

[1] see Bernanke and Boivin [4] and Bernanke, Boivin, and Eliasz [6] for a discussion.

this approach is that one only need to identify those shocks of interest[2] and not the whole system. This makes the approach easy to implement and appealing. Hence the estimated DSGE does not need to be fully stochastically specified. The second approach builds on the idea to derive restrictions from the DSGE models for both the estimation of the parameters of the VAR model and the identification of the model (see Del Negro and Schorfheide [12] and Sims [2008]). Del Negro and Schorfheide [12] provide a very useful tool for forecasting and policy analysis by exploiting an estimated DSGE model as a source of a prior to estimate a VAR resulting in a DSGE-VAR. This tool allows evaluating the validity of the DSGE model, produces improved forecasts and allows improved policy analysis. I suggest extending this approach in several dimensions. The estimation of the dummy observations can be based on a standard indicator-based[3] DSGE model or a Factor-based DSGE model as proposed by Boivin and Giannoni. The latter approach has been shown to produce reduced measurement error and improved forecasts. Del Negro and Schorfheide estimate their model several times with different fixed values for the weight and hence importance assigned to the DSGE model and chose the one that maximizes the marginal data density. Opposed to them I treat this weighting parameter as random and estimate it as part of the DSGE parameter space hence averaging over it. This approach is more efficient from a computational point of view and is more Bayesian in spirit. This weighting parameter is an interesting metric to evaluate the potential misspecification of the DSGE models and furthermore indirectly provide a new class of structural Bayesian VAR or DFM. Here the use the DSGE model to shape the prior odds for a dynamic factor model and provide an identification scheme consistent with the theoretical models. The optimal weight on the DSGE model gives a new dimension to access the relevance of the economic restrictions implied by the theoretical model. The proposed approach gives insight to what extend models are good in fit, their implied tight zero restriction are valid and allows to access the degree of misspecification and biased inference based on single indicators to represent economic concepts.

I find a high degree of comovement across the countries on the real side if the economy. Regarding the inflation indicator a higher degree of comovement can be found considering year-on-year inflation and a low degree of comovement for the quarter-on-quarter specification. The monetary transmission mechanism turns out to be mostly homogenous though some differences appear in the scale of the reaction of the labor market. Real wages and employment in France and Germany show a lower exposure to a contractionary monetary policy compared to the euro are common factors whereas Finland, Portugal and Spain react 5-10 times stronger in scale. Monetary policy shocks have a constant contribution of around 8% of the variation of Business cycle components over the entire horizon considered. The key drivers are productivity and investment shocks that sum to a constant share of around 75% of the variation in the business cycle components. Interestingly productivity shocks increase their relative contribution over the horizon whereas investment shocks decries in relative contribution by the same amount. The sign restriction approach within the DSGE-DFM framework produces less uncertainty in the impulse response functions than the DSGE rotation

[2]see Uhlig (2005).

[3]I refer to "*indicator-based*" DSGE model when only single time series are chosen to correspond to economic concepts formulated in by the DSGE model to be estimated.

identification in the DSGE-DFM framework. These two approaches are restrictive in the sense that the transmission mechanism allows only heterogeneity regarding the scale of responses. For richer and more complicated dynamics in the transmission mechanism I provide results in the DFM framework combined with robust sign restriction derived from economic theory via an estimated DSGE model. The flexibility comes at the price of more sign restrictions to be imposed.

4.2 Framework

In order to analyze the degree of comovement and the transmission mechanism of macroeconomic shocks in the Euro area I will estimate three different econometric models all based on the same large set of Euro area data. The first model is a Bayesian dynamic factor model, second I will estimate a factor-based DSGE model in the spirit of Boivin and Giannoni which nests the indicator based DSGE estimation which is the standard approach in the literature (see e.g. among others Smets and Wouters. Finally I will explain the proposed DSGE-DFM model. The resulting DSGE-DFM model I propose can be considered as a factor-generalization of their approach.

4.2.1 Bayesian Dynamic Factor Analysis

The dynamic factor model considered here is cast in the state space form and jointly estimated via MCMC methods to get the posterior distribution of the model parameters and states. The superscript E is attached for the DFM model to distinguish between the three models. Consider

$$F_t^E = \sum_{p=1}^{P} B_p^E F_{t-p}^E + u_t \quad \text{with} \quad u_t \sim iidN\left(0, \Sigma^E\right) \tag{4.2.1}$$

$$X_t = \Lambda^E F_t^E + e_t^E \tag{4.2.2}$$

$$e_{nt}^E = \sum_{q=1}^{Q} \psi_{nq}^E e_{nt-q}^E + \varepsilon_{nt} \quad \text{with} \quad \varepsilon_{nt}^E \sim iidN\left(0, R_{nn}^E\right) \tag{4.2.3}$$

where $n = 1, ..., N$, $t = Q + 1, ..., T$, X_t is an $N \times 1$ data vector, F_t^E is a $K \times 1$ vector of unobservable factors, Λ^e is an $N \times K$ matrix of factor loadings, $e_t^E = \left(e_{1t}^E, ..., e_{Nt}^E\right)'$ is an $N \times 1$ vector of unobservable idiosyncratic components, u_t is a $K \times 1$ vector of factor innovations, and B_p^e is a $K \times K$ matrix for each $p = 1, ..., P$. The variables e_{nt} are assumed to follow pairwise independent Gaussian autoregressive processes of order Q. Let x_{nt} denote the n'th observable variable in period t, for each $n = 1, ..., N$. Let f_{kt}^E denote the k'th unobservable factor in period t, for each $k = 1, ..., K$. We have $X_t = (x_{1t}, ..., x_{Nt})'$ and $F_t^E = \left(f_{1t}^E, ..., f_{Kt}^E\right)'$. Furthermore note that I will use the superscript E to denote variables and parameters referring to the *p*ure empirical DFM to be distinguished from the F-DSGE which will be denoted by F^M and the DSGE-DFM without any superscript.

4.2.2 Factor-based DSGE estimation

This section draws on the approach employed by Boivin and Giannoni in a slightly simplified version[4]. The key motivation for the factor-based as opposed to the single indicator estimation of DSGE models follows and builds on the rich empirical evidence of dynamic factor analysis. The parsimonious representation of large information sets crucial for describing the evolution of pertinent macroeconomic variables and economic concepts researchers have in mind, reflecting realistically the information set agents and policy makers have prior to taking decision, reduced measurement error and avoiding omitted variable bias.

Employing the approach suggested by Sims the general linear (or linearized) rational expectations model is of the following form[5]:

$$\Gamma_0(\gamma)S_{t+1} = \Gamma_1(\gamma)S_t + \Psi(\gamma)\varepsilon_t + \Pi(\gamma)\eta_t$$

where the vector $S_t \equiv [Z_t', s_t']'$ denotes the vector of state variables, Z_t denotes predetermined endogenous or lagged exogenous variables and s_t denotes exogenous variables following a stochastic process, ε_t is the vector of exogenous shocks and η_t is the vector collecting the expectational errors implying the respective equation to be added to the system. The fundamental solution to the DSGE model is given by

$$S_t = G(\gamma)S_{t-1} + H(\gamma)\varepsilon_t$$

Here the matrices $G(\gamma)$ and $H(\gamma)$ are the highly nonlinear functions of the parameters once the model is solved. Furthermore the relation of the states S_t to the vector containing the non-predetermined endogenous variables z_t is expressed through

$$z_t = DS_t. \tag{4.2.4}$$

The equation (4.2.4) refers to the standard indicator variables based DSGE estimation where for each economic concept there is one time time-series to be referred to in the evaluation procedure of the likelihood. The potentially large number of variables of interest in the factor-based DSGE estimation procedure of Boivin and Giannoni are collected in the vector F_t and define a linear combination to the state variables of the following form:

$${F^M}_t = \Phi S_t$$

where Φ depends on the model parameters and the selection of the variables in F^M.

[4]I only focus on the approach where the data is blocked according to the respective economic concept that the DSGE model is supposed to describe. The respective factors are extracted from the grouped blocks of data according to the economic concepts. No further data block is considered to improve the parameter estimation as it is proposed by Boivin and Giannoni for simplification purposes. In this paper I focus rather on the identification schemes and the structural factors. For a more detailed discussion about the approach the interested reader can find a detailed description in the original paper.

[5]Here I employ the solution method of Sims. Alternative solution methods such as proposed by Uhlig work equally well and implementation are readily available through Uhlig's toolkit.

The final state space representation of the DSGE model is:

$$X_t = \Lambda^M S_t + e_t^M \tag{4.2.5}$$
$$S_t = G(\gamma) S_{t-1} + H(\gamma)\varepsilon_t \tag{4.2.6}$$
$$e_t^M = \Psi^M e_{t-1}^M + \varepsilon_t^M, \quad \varepsilon_t^M \sim i.i.d. \mathcal{N}(0, R^M) \tag{4.2.7}$$

where

$$\Lambda^M = \Lambda_F \Phi \quad \text{and} \quad {F^M}_t = \Phi S_t$$

Here Λ_F implies the highly structured normalization in order to get the structural rotation for the factors of interest. Normalization issues are discussed more elaborately in section (4.2.4).

4.2.3 DSGE-DFM Approach

The proposed DSGE-DFM model combines the approaches of the previous two explained. Strictly speaking I estimate a dynamic factor model as in section (4.2.1) which includes at each stage of the iteration the dummy observations approach based on a modified likelihood function which is also a function of the deep parameters and implied moments of the estimated DSGE model in section (4.2.2). Intuitively this amount to take the estimated DSGE model as a dummy observation prior for the estimation of the DFM. This results in the DSGE-DFM approach that I will describe in this section. Note that this approach can be considered as a factor generalization of the DSGE-VAR approach introduced by Del Negro and Schorfheide [12]. The companion form of the state space representation of the DFM in (4.2.1) - (4.2.3) is given by:

$$X = F\Lambda' + E \tag{4.2.8}$$
$$F = ZB' + U \tag{4.2.9}$$

where

$$X = \begin{bmatrix} X_1' \\ X_2' \\ \vdots \\ X_T' \end{bmatrix}, \quad Z = \begin{bmatrix} {F_P}' & {F_{P-1}}' & \cdots & {F_1}' \\ F_{P+1}' & F_P' & \cdots & F_2' \\ \vdots & \vdots & \ddots & \vdots \\ F_{T-1}' & F_{T-1}' & \cdots & F_{T-P}' \end{bmatrix}, \quad F = \begin{bmatrix} F_{P+1}' \\ F_{P+2}' \\ \vdots \\ F_T' \end{bmatrix}.$$

Note that this representation is very close to the state space representation in subsection (4.2.1), however is it important to note that here the factors $F = f(\gamma, \lambda, X^T, B_p, \Sigma_u)$ are also a function of the DSGE model parameters. The estimation of factors invokes the additional estimation of a DSGE model to deliver the posterior of the deep parameters on which F_t also depends unlike the case of a "pure" dynamic factor model in (4.2.1). This shall be explained in greater detail in the following sections.

4.2.4 Normalization of the factors

In DFM it is of great importance to impose further restrictions for the model parameters and the factors to be uniquely identified against rotational, scale and sign indeterminacy. This task is of crucial importance as the likelihood is salient about the specific unique rotation that separately identifies factors, loadings and parameters. Different rotations are observationally equivalent resulting with the same likelihood although the model can be very different. There are different assumptions one can impose depending on the purpose of the analysis. One standard approach to the identification of factor models is the approach goes back to Geweke and Zhou. They restrict the upper $K \times K$ block of the factor loading matrix Λ to be lower triangular. Alternatively one could impose the upper $K \times K$ to be identity. Note that this restriction is over-identified however very convenient as no further restrictions are required for the scale and sign determinacy. A third alternative is to group the data according to some model the researcher has in mind and extract the respective factors solely from these predefined blocks of data. This approach has implied restrictions on the factor loading matrix resulting in over-identification. However the advantage is that it allows to label the factor according to some model the researcher has in mind. Let's summarize the three approaches to identification her

1. Lower triangular Block in Λ
2. Identity Block in Λ
3. Block-diagonally of Λ according to specific model

In this paper I employ the second (Identity Block) and the third version (Block-diagonally). The block-diagonally of the factor loading matrix Λ furthermore requires to set in each block the first element to 1 as a normalization and I impose all elements to be positive. I block the data according to respective economic concepts the DSGE model is supposed to reflect. Note that the grouping of the data and the block-diagonality assumptions are additional restrictions imposed for the factors to represent solely the economic concepts for the DSGE model. Hence the structure of the factor loading imposed combines the normalization and the additional restrictions. For unique identification of factors and loadings it is sufficient to set the upper $[K \times K]$ block to identity and no further restrictions are required. But then there is no interpretation of the factors as economic concepts which is not what we want. For that the additional assumptions of block diagonality according to the respective groups of data is required. This way the extracted factors are consistent with the respective economic concept as each data block only loads with one of the corresponding factors. Here a general example how the factor loading matrix looks like with the described restriction

$$\Lambda = \begin{bmatrix} \Lambda_{F,1} & 0 & \cdots & 0 \\ 0 & \Lambda_{F,2} & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \Lambda_{F,K} \end{bmatrix} \text{ with } \quad \Lambda_{F,k} = \begin{bmatrix} 1 \\ \lambda_{F,k}^{2} \\ \lambda_{F,k}^{3} \\ \vdots \\ \lambda_{F,k}^{n_k} \end{bmatrix} \tag{4.2.10}$$

where in each block of factor loadings $\Lambda_{F,k}$ with $k = 1, \ldots, K$ the first element is set to 1.[6]

4.3 Estimation and Inference

The estimation algorithm of all three models (i) DFM, (ii) Factor-based DSGE and (iii) DSGE-DFM described previously are based on Markov chain Monte Carlo (MCMC) methods to simulate the posterior distribution of the unknown parameters and latent states of interest. The reason is that in all three cases the parameter space to be estimated is of large dimension so that due to the curse of dimensionality estimation via maximum likelihood is infeasible. MCMC methods are becoming increasingly popular as an estimation device to simulate the joint posterior density of large parameter spaces. These methods allow to break down a joint density that might be hard or impossible to sample from into conditional ones one can readily sample from to approximate the target distribution. The Gibbs sampler is a special method of the class of MCMC algorithms. It is an iterative procedure to sample from conditional distributions. Strictly speaking for model (i) I employ a Gibbs sampler and for model (ii) and (iii) I employ a hybrid version namely a Metropolis-within-Gibbs sampler.

4.3.1 Sampling Algorithm 1: Dynamic Factor Model

The estimation of Bayesian dynamic factor models is facilitated via multi-move Gibb sampling that goes back to Carter and Kohn [10], Frühwirth-Schnatter [16]. We seek to get the posterior distribution of the parameter space $\Theta^E = (\Lambda^E, R^E, \Psi^E, B^E, \Sigma^E)$ and the history of unobserved factors $F^{T(E)}$. Note that opposed to the following models here we do not make use of the DSGE model at all. The approach of the multi-move Gibbs sampler involves to sample the parameter blocks and the unobserved factors visa the Kalman smoother. The parameters can be sampled from the conditional blocks as give the data and the unobserved factors the parameters of the observation equation and the state equation can be sampled from independently. A pseudo code for the estimation procedure to estimate the dynamic factor model can be found in algorithm (4.1). Note that for ease of notation the superscripts E and M describing to which model they refer are omitted. However it is clear from the caption.

Step 1: Sample $p(F^{T(E)} \mid \Theta^E, X^T)$. Follows exactly the previous subsection. But note that there is no conditioning on the deep parameters γ of the DSGE model.

Step 2: Sample $p(\Lambda^E, R^E \mid F^{(E)T}, \Psi^{E(g-1)}, X^T) p(\Psi^E \mid F^{(E)T}, \Lambda^{E(g)}, R^{E(g)}, X^T)$. But note that there is no conditioning on the deep parameters of the DSGE model. The respective distributions follow

[6] See Belviso and Milani [3] for a Structural FAVAR approach.

Prior Specification.

$$\begin{aligned} R^E_{nn} &\sim \mathcal{IG}(\delta_0/2, \nu_0/2) \\ \Lambda^E_n &\sim \mathcal{N}(\Lambda_0, {R^E_{nn}}^{-1} M_0^{-1}) \\ \Psi^E_{kk} &\sim \mathcal{N}(0,1) \end{aligned}$$

where the degrees of freedom is $\nu_0 = 0.001$ the prior scale is $\delta_0 = 6$, the prior variance in the coefficient of each observation equation is $M_0 = I_K$ and the prior mean of the loadings is $\Lambda_0 = 0$. The prior for the coefficients of the idiosyncratic process Ψ_{kk} are standard normal.

Step 2: Sample $p(B^E, \Sigma^E \mid F^{(E)T}, X^T)$. But note that there is no conditioning on the deep parameters of the DSGE model. To draw $vec(B^E)$ and Σ^E conditional on the most current draws of the factors and the data I employ the Normal-Wishart prior according to Uhlig [1994]

$$\begin{aligned} p(\Sigma^E) &= \mathcal{IW}(S_0, \nu_0) \\ p(vec(B^E) \mid \Sigma^E) &= \mathcal{N}(\bar{B^E}_0, \Sigma^E \otimes N_0^{-1}) \end{aligned}$$

with prior specification

$$\begin{aligned} \nu_0 &= K + 2 \\ N_0 &= 0_{K \times K} \end{aligned}$$

where the choice of S_0 and $\bar{B}_0$ are arbitrary as they cancel out in the posterior. To ensure stationarity, we truncate the draws by discarding the draws of B^E with the larges eigenvalue greater than 1 in absolute value.

4.3.2 Sampling Algorithm 2: Factor-based DSGE estimation

This section draws on the paper of Boivin and Giannoni [2006] who first combined the estimation of DSGE model with a larger set of indicator variables, hence combining multiple informative indicators with the DSGE estimation. Their idea is to combine the advantages of large dimensional factor models for a more accurate estimation of the models concepts and shocks. They show that their approach avoids the omitted variable bias and implies different conclusions about key structural parameters and sources of economic fluctuations. This subsection briefly explains the estimation procedure and the single steps required to do inference on the deep parameters of the DSGE model and the factors via MCMC methods. For a more elaborative description of the framework and the estimation procedure see the original paper. I provide a detailed description of the estimation procedure in Appendix (4.7). Algorithm (4.2) provides a pseudo code of this approach to summarize the key steps to do inference. Note that in Algorithm (4.2) the superscript M is omitted for ease of notation.

First lets define γ as the deep parameter vector of the solution of the DSGE model to be estimated individually. Furthermore let $\theta_M = (\Lambda^M, R^M, \Psi^M)$ be the parameter space of the observation equation that are independent of the ones from the state equation.

Table 4.1: **Algorithm 2**

Algorithm 2: DFM estimation via Multi-move Gibbs sampling.

Step 0, [Initialization]: $p_0(F^0, \lambda^{f0}, \lambda^{y0}, b^0, R_e^0, Q_u^0)$.
`Set` $g \rightsquigarrow 0$.
`Get initial values for states and parameters.`
`Set` $g \rightsquigarrow 1$.

Step 1, [Evaluate likelihood of latent states]: $p(F^T \mid X^T, \Lambda, \Psi, R, B, \Sigma) \sim FFBS$
`Do forward filtering and backward sampling (FFBS)`

Step 2, [Sample parameters from observation equation]: $p(\Lambda_n, \Psi_{nn}, R_{nn} \mid X^T, F^T)$
$2.a: \quad p(\Lambda_n, R_{nn} \mid \Psi_{nn}^{(g-1)}, X^T, F^T) \sim f_{\mathcal{N}-\mathcal{IG}}$
$2.b: \quad p(\Psi_{nn} \mid \Lambda_n^{(g)}, R_{nn}^{(g)}, X^T, F^T) \sim f_{\mathcal{N}}$
`Sample equation by equation (conditional Gaussianity).`

Step 3, [Sample parameters from state equation]: $p(B, \Sigma \mid X^T, F^T)$
$3.a: \quad p(\Sigma \mid F^{T(g)}, X^T) \sim f_{\mathcal{IW}}$
$3.b: \quad p(B \mid \Sigma^{(g)}, F^{T(g)}, X^T) \sim f_{\mathcal{N}}$
`Sample from a normal inverted Wishart density.`
`If` $g \leq G$ `set` $g \rightsquigarrow g+1$ `and go to Step 1.`
`Otherwise stop.`

The task is to sample the joint estimation of the parameters of the state space system and the unobserved latent states $p(\Theta, S^T)$ where $\Theta = (\gamma, \theta_M)$. We have to iteratively sample through the following three conditional distributions:

$$\text{Step 1:} \quad p(F^{T(g)} \mid \gamma^{(g-1)}, \theta_M^{(g-1)}, X^T) \tag{4.3.1}$$

$$\text{Step 2:} \quad p(\Lambda^M, R^M \mid F^T, \Psi^{M(g-1)}, X^T) p(\Psi^M \mid F^T, \Lambda^{M(g)}, R^{M(g)}, X^T) \tag{4.3.2}$$

$$\text{Step 3:} \quad p(B^{M(g)}, \Sigma^{M(g)}, \gamma \mid F^T) \tag{4.3.3}$$

4.3.3 Sampling Algorithm 3: DSGE-DFM

In this subsection I will very briefly explain how to merge the previous two approaches to get the DSGE-DFM model. Here I will lay down how exploit an estimated factor-based DSGE model as a prior to estimate a DFM. One motivation is to exploit the large data set for both, the dynamic factor model and the DSGE model to serve as a prior. Thus omitted variable bias and a more accurate estimation of the model concepts and shocks are facilitated. Furthermore this approach offers a way to assign an interpreta-

Table 4.2: **Algorithm 3**

Algorithm 3: Factor-Based DSGE Estimation.

Step 0, [Initialization]: $p_0(F^0, \lambda^{f0}, \lambda^{y0}, b^0, R_e^0, Q_u^0)$.

Set $g \rightsquigarrow 0$.
Get initial values for states and parameters.
Set $g \rightsquigarrow 1$.

Step 0, Initialization: $p(\gamma)$.

Set $g \rightsquigarrow 0$.
Get initial values for γ.
Set $g \rightsquigarrow 1$.

Step 1, DSGE model part:.

Step 1.1, Candidate draw: $p(\gamma)$.

Draw candidate from proposal distribution$\gamma\hat{}*$.

Step 1.2, Solving the model:

Solve model for candidate draw. Set up state space.

Step 1.3, Evaluate proposal:

Evaluate prior $p(\gamma\hat{}*)$ and $L(X\hat{}T, \gamma\hat{}*)$.

Step 1.4, Accept/Reject:

Draw $prob \sim U(0,1)$.
Accept with probability $prob \leq \frac{L(X\hat{}T;\gamma\_c)p(\gamma\_c)}{L(X\hat{}T;\gamma\hat{}(g-1))p(\gamma\hat{}(g-1))}$ and set $\gamma\hat{}g = \gamma\_c$
otherwise reject and set $\gamma\hat{}g = \gamma\_(g-1)$

Step 2, Evaluate Likelihood of states S or F: $p(F^T \mid X^T, \Lambda, \Psi, R, B, Q, \gamma) \sim FFBS$

Do forward filtering and backward sampling

Step 3, Observation Equation: $p(\Lambda, R \mid X^T, F^T, \Psi) \sim NIG$ and $p(\Psi \mid X^T, F^T, \Lambda, R) \sim N$

Posterior sampling of parameters in observation equation.
$R\_nn \sim IG\_2(3, .001)$
$\Lambda\_n \sim N(\Lambda\_0, V\_\lambda)$
$\Psi \sim N(\Psi\_0, V\_\Psi)$

tion to the factors based on economic theory via the DSGE model in addition to the normalization and additional restrictions imposed on the factor loading matrix. We want

Table 4.3: **Algorithm 4**

Step 0, Initialization: $p_0(F^0, \Lambda^0, \Psi^0, R^0, B^0, \Sigma^0, \gamma^0, \lambda^0)$.

```
Get initial values and set g ⇝ 1.
```

Step 1, Sample factors via Kalman filter: $p(F^T \mid X^T, \Lambda, \Psi, R, B, \Sigma, \gamma, \lambda) \sim FFBS$

```
Do forward filtering and backward sampling
```

Step 2, Sample parameters from observation equation: $p(\Lambda_n, \Psi_{nn}, R_{nn} \mid X^T, F^T) \sim N - IG$

```
Sample parameters equation by equation
from a normal inverse gamma density.
```

Step 3, Sample parameters from state equation: p$B, \Sigma, \gamma, \lambda \mid X^T, F^T$

```
Invoke Algorithm 2.
If g ≤ G set g ⇝ g+1 and go to Step 1.
Otherwise stop.
```

to do inference on the factors F^T and the parameter space $\Theta = (\Lambda, R, \Psi, B, \Sigma, \gamma, \lambda)$. The joint density can be blocked in the following form

$$
\begin{aligned}
p(F^T, \Lambda, R, \Psi, B, \Sigma, \gamma, \lambda) =& p(F^T \mid \Lambda, R, \Psi, B, \Sigma, \gamma, \lambda, X^T) \\
& \times p(\Lambda, R, \Psi \mid F^T, X^T) \\
& \times p(B, \Sigma, \gamma, \lambda, \mid \Lambda, R, \Psi, B, F^T, X^T)
\end{aligned}
$$

resulting in three main blocks one has to sample from iteratively. The estimation procedure is exactly like the one for the DFM except that an additional step has to be included to invoke the Del Negro and Schorfheide approach when sampling the state equation parameters, hyperparameters and the deep parameters of the DSGE model. The approach can be summarized by the following three main steps at each iteration g

$$
\begin{aligned}
&\text{Step 1:} \quad p(F^{T(g)} \mid \Lambda^{(g-1)}, R^{(g-1)}, \Psi^{(g-1)}, B^{(g-1)}, \Sigma^{(g-1)}, \gamma^{(g-1)}, X^T) \\
&\text{Step 2:} \quad p(\Lambda^{(g)}, R^{(g)}, \Psi^{(g)} \mid F^{T(g)}, X^T) \\
&\text{Step 3:} \quad p(B^{(g)}, \Sigma^{(g)}, \gamma^{(g)}, \lambda^{(g)} \mid \Lambda^{(g)}, R^{(g)}, \Psi^{(g)}, F^{T(g)}, X^T).
\end{aligned}
$$

The crucial innovation of the approach lies in **S**tep 3 where the approach of Del Negro and Schorfheide [12] is invoked. In the next section I will explain the respective steps for the estimation.

4.3.3.1 Step 1: Sample $p(F^T \mid \gamma^{(g-1)}, \theta_F^{(g-1)}, X^T)$

Sampling the factors follows the multi move Gibbs sampling approach by Carter and Cohn (1994), Frühwirth-Schnatter (1994) surveyed in Kim and Nelson to sample the history of the factors F^T as the product of conditional distributions at each date t as follows

$$p(F^T \mid \gamma,, X^T) = p(F_T \mid \gamma, \theta_F, X^T) \prod_{t=1}^{T} p(F_t \mid F_{t+1}, \gamma, X^T)$$

which relies on the Markov property of the F_t. Given that the state space representation in linear and Gaussian, we have

$$\begin{aligned} F_T \mid \gamma, \theta_F, X^T &\sim \mathcal{N}(F_{T|T}, P_{T|T}) \\ F_t \mid F_{t+1}, \gamma, \theta_F &\sim \mathcal{N}(F_{t|t,F_{t+1}}, P_{t|t,F_{t+1}}) \end{aligned}$$

where

$$\begin{aligned} F_{T|T} &= E[F_T \mid \gamma, \theta_F, X^T] \\ F_{T|T} &= cov(F_T \mid \gamma, \theta_F, X^T) \\ F_{t|t,F_{t+1}} &= E[F_T \mid F_{t+1}, \gamma, \theta_F, X^T] \\ &= F_{t|t} + P_{t|t} B'(G P_{t|t} G' + \Sigma)^{-1}(F_{t+1} - B F_{t|t}) \\ P_{t|t,F_{t+1}} &= cov(F_T \mid F_{t+1}, \gamma, \theta_F, X^T) \\ &= P_{t|t} + P_{t|t} B'(B P_{t|t} B' + \Sigma)^{-1} B P_{t|t}, \end{aligned}$$

here the standard notation holds as in Kim and Nelson (1999) for the expectation conditional on the respective information set.

4.3.3.2 Step 2: $p(\Lambda, R \mid F^T, \Psi^{(g-1)}, X^T) p(\Psi \mid F^T, \Lambda^{(g)}, R^{(g)}, X^T)$

Here we have to take care of the autocorrelation in the observation equation due to the autoregressive process in the idiosyncratic process. The following quasi-differenced equations have to be introduced

$$\begin{aligned} \tilde{X}_t &= X_t - \Psi X_{t-1} \\ \tilde{F}_t &= F_t - \Psi F_{t-1} \\ \tilde{X}_t &= \Lambda \tilde{F}_t + \nu_t. \end{aligned}$$

Note that R is diagonal and hence given F^T and Ψ standard regression methods can be applied to do inference equation by equation. Here I follow closely Boivin and Giannoni (2007) by setting the prior for each equation as

$$\begin{aligned} R_{nn} &\sim \mathcal{IG}(\delta_0/2, \nu_0/2) \\ \Lambda_n &\sim \mathcal{N}(\Lambda_0, R_{nn}^{-1} M_0^{-1}) \end{aligned}$$

where the degrees of freedom is $\nu_0 = 0.001$ the prior scale is $\delta_0 = 6$ and the prior variance in the coefficient of each observation equation is $M_0 = I_K$. The posterior distribution is given by

$$\begin{aligned} R_{nn} \mid F^T, X^T \sim & \mathcal{IG}(\delta_T, \nu_T) \\ \Lambda_n \mid F^T, X^T \sim & \mathcal{N}(\bar{\Lambda}_n, R_{nn}^{-1} \bar{M}_0^{-1}) \end{aligned}$$

where

$$\begin{aligned} \nu_T =& T + \nu_0 \\ \delta_T =& \delta_0 + u'u + (\Lambda_n - \hat{\Lambda}_n)'(M_0^{-1} + (Z'Z)^{-1})^{-1}(\Lambda_n - \hat{\Lambda}_n) \\ \bar{M}_n =& M_0 + (Z'Z)^{-1} \end{aligned}$$

Give the previous draws we have $\nu_t = X_t - \Lambda F_t$ and assuming a standard normal prior

$$\Psi_{kk} \sim \mathcal{N}(0, 1)$$

the posterior is

$$\Psi_{kk} \mid F^T, \Lambda^{(g)}, R^{(g)}, X^T \sim \mathcal{N}(\bar{\Psi}_k k, \bar{N}^{-1})$$

where

$$\begin{aligned} \bar{\Psi}_{kk} =& \bar{N}^{-1} R_{kk}^{-1} \hat{\nu}_n' \hat{\nu}_n \hat{\Psi}_k k \\ \bar{N}^{-1} =& 1 + R_{kk}^{-1} \hat{\nu}_n' \hat{\nu}_n \end{aligned}$$

where $\hat{\Psi}_k k$ denotes the OLS estimate of the residuals on its lagged values.

4.3.3.3 Step 3: $p(B, \Sigma, \gamma, \lambda, \mid \Lambda, R, \Psi, B, F^T, X^T)$

Here the key part to explain the estimation part for the DSGE-DFM part starts. The hierarchical prior to conduct Bayesian inference for the state equation is of the form

$$p(B, \Sigma, \gamma, \lambda) = p(B, \Sigma \mid \gamma, \lambda) p(\gamma) p(\lambda) \tag{4.3.4}$$

where γ is the vector of DSGE model parameters and λ is the weight assigned to the reliability of the DSGE model relative to the actual data length.

Step 3.1: Sample $p(\gamma \mid F^T)$. Let γ^0 be the posterior mode, Σ_γ the inverse of the Hessian computed at the posterior mode, c the scaling factor and g the current iteration step. In order to draw the DSGE parameters from the posterior distribution one has to go through the following steps:

1. Choose a starting point for γ^0
2. Draw a proposal γ^* from a jumping distribution $J(\gamma^* \mid \gamma^{g-1} = \mathcal{N}(\gamma^{g-1}, c\Sigma_\gamma))$
3. Compute acceptance ratio $\alpha = \frac{p(F^{T(g)} \mid \gamma^*) p(\gamma^*)}{p(F^{T(g)} \mid \gamma^{g-1}) p(\gamma^{g-1})}$
4. Accept the proposal draw $\gamma^g = \gamma^*$ with probability $min(\alpha, 1)$ otherwise reject and keep previous draw and set $\gamma^g = \gamma^{g-1}$.

In order to compute the acceptance ratio the likelihood of the factors given the parameters $p(F^{T(g)} \mid \gamma)$ one has to evaluate the following likelihood[7]

$$p_\lambda(F^T \mid \gamma) = \frac{|\lambda T \Gamma_{FF}(\gamma) + F'F|^{-\frac{n}{2}} \left|(1+\lambda)T\hat{\Sigma}_b(\gamma)\right|^{-\frac{(1+\lambda)T-k}{2}}}{|\lambda T \Gamma_{FF}(\gamma)|^{-\frac{n}{2}} \, |\lambda T \Sigma * (\gamma)|^{-\frac{\lambda T - k}{2}}} \times \frac{(2\pi)^{-\frac{nT}{2}} 2^{\frac{n((1+\lambda)T-k)}{2}} \prod_{i=1}^n \Gamma[((1+\lambda)T-k+1-i)/2]}{2^{\frac{n(\lambda T-k)}{2}} \prod_{i=1}^n \Gamma[(\lambda T - k + 1 - i)/2]}. \tag{4.3.5}$$

Step 3.2: Sample $p(B, \Sigma \mid \gamma, F^T)$. The dummy observation prior based on the estimated DSGE model means intuitively to augmenting the actual data which in our case is the sample of the extracted structural factors F with artificial dummy observations of length $T^* = \lambda T$. The idea of artificial dummy observation to be added to the actual data for inference dates back to Theil and Goldberger [1961] and most prominently advocated by Sims [2005] and Sims and Zha [1998] in the context of Bayesian VARs. For an elaborate discussion please refer to the paper by Del Negro and Schorfheide [12]. It is important to note that for the final implementation we actually do not generate data from the estimated DSGE model. Generating artificial data from the recursive law of motion of the solved DSGE model conditional on the sampled γ, augment it with the actual sampled factors and combine it with the likelihood results in

$$p(B, \Sigma \mid \gamma, \lambda) = c^{-1}(\gamma, \lambda) \, |\Sigma|^{-\frac{(\lambda T + K)}{2}} \times e^{-\frac{1}{2} tr\left[\Sigma^{-1}(F^{*\prime}F^* - B'Z^*F^* - BZ^{*\prime}F^{*\prime} + B'Z^{*\prime}Z^*B)\right]}. \tag{4.3.6}$$

Here (F^*, Z^*) refer to the simulated dummy observations respectively. As noted before we don't want to use the simulated dummy observations hence the idea is to replace the artificial DSGE model generated sample moments $[F^{*\prime}F^*, F^{*\prime}Z^*, Z^{*\prime}Z^*]$ by the theoretical population analogues $[\Gamma^*_{FF}(\gamma), \Gamma^*_{ZZ}(\gamma), \Gamma^*_{ZF}(\gamma)]$ conditional on (γ, λ) which yields the prior likelihood function of the form

$$p(B, \Sigma \mid \gamma, \lambda) = c^{-1}(\gamma, \lambda) \, |\Sigma|^{-\frac{\lambda T + K}{2}} \times e^{-\frac{1}{2} tr\left[\lambda T \Sigma^{-1}(\lambda T \Gamma^*_{FF}(\gamma) - B'\Gamma^*_{ZF}(\gamma) - \Gamma^*_{FZ}(\gamma)B + B'^*\Gamma^*_{ZZ}(\gamma)B)\right]} \tag{4.3.7}$$

Note that a proper prior requires that $\lambda T \geq K(P+1)$ for it to be proper. The proportionally factor $c(\gamma, \lambda)$ ensures that the density integrates to one. For a more elaborate derivation please refer to the Appendix of DS (2004). The resulting maximum likelihood estimators is

$$vec(B^*(\gamma)) = \Gamma_{ZZ}(\gamma)^{-1}\Gamma_{ZF}(\gamma) \tag{4.3.8}$$

$$\Sigma^*(\gamma) = \Gamma_{FF}(\gamma) - \Gamma_{FZ}(\gamma)\Gamma_{ZZ}(\gamma)^{-1}\Gamma_{ZF}(\gamma). \tag{4.3.9}$$

Conditional on γ the prior distribution of the DFM parameters is of the Inverted-Wishart-Normal form :

$$\Sigma \mid \gamma, \lambda \sim \mathcal{IW}(\lambda T \Sigma^*(\gamma), \lambda T - KP - K) \tag{4.3.10}$$

$$vec(B) \mid \Sigma, \gamma, \lambda \sim \mathcal{N}\left(vec(B^*(\gamma)), \Sigma \otimes [\lambda T \Gamma_{ZZ}(\gamma)]^{-1}\right) \tag{4.3.11}$$

[7] For a derivation please refer to Del Negro and Schorfheide

note that this approach requires $\Gamma_{ZZ}(\gamma)$ to be invertible hence the respective DSGE model generating the artificial dummy observations must be stochastically fully defined which means that the model has to have as many shocks as observables. One way to easily avoid the singularity of $\Gamma_{ZZ}(\gamma)$ is to introduce measurement error.[8].

Posterior Derivation. The posterior distribution may be factorized the following way

$$p(B, \Sigma, \gamma, \lambda \mid F^T) = p(B, \Sigma \mid F^T, \gamma, \lambda) \times p(\gamma, \lambda \mid F^T) \tag{4.3.12}$$

where F^T denotes the extracted structural sample conditional on γ . Note that it can be very well also a non structural sample of the factors. But for the empirical application at hand I employ a structural DSGE-DFM approach. The first term on the right hand side of the above expression has a closed form expression which conditional on γ and λ is a conjugate prior. This means that conditional on $(\Sigma, \gamma, \lambda, F^T)$ the distribution of B is matric-variate normal and conditional on $(\gamma, \lambda, \mathcal{F}^T)$ the distribution of Σ is inverted Wishart which result in the following expressions

$$vec(B) \mid \Sigma, \gamma, \lambda, F^T \sim \mathcal{N}\left(vec(\widetilde{B}(\gamma, \lambda)), \Sigma \otimes N(\gamma, \lambda)^{-1}\right) \tag{4.3.13}$$

$$\Sigma \mid \gamma, \lambda, F^T \sim \mathcal{IW}\left((1+\lambda)T\widetilde{\Sigma}(\gamma, \lambda), (1+\lambda)T - KP - K\right) \tag{4.3.14}$$

where

$$\widetilde{B}(\gamma, \lambda) = (\lambda T \Gamma_{ZZ}(\gamma) + Z'Z)^{-1}(\lambda T \Gamma_{ZF}(\gamma) + Z'F) \tag{4.3.15}$$

$$\widetilde{\Sigma}(\gamma, \lambda) = \frac{1}{(1+\lambda)T}[(\lambda T \Gamma_{FF}(\gamma) + F'F) \tag{4.3.16}$$

$$- (\lambda T \Gamma_{FZ}(\gamma) + Y'Z)(\lambda T \Gamma_{ZZ}(\gamma) + Z'Z)^{-1}(\lambda T \Gamma_{ZF}(\gamma) + Z'F)]. \tag{4.3.17}$$

As there is no closed for expression for the joint posterior density of (γ, λ) (which is the second term) so as is standard in the literature I employ a MCMC algorithm to get the posterior. Note the difference compared to Del Negro and Schofheide [2004] in that the prior weight λ is a random parameter and part of the parameter vector γ however note the orthogonality between the two: $\gamma \perp \lambda$.

Digression on estimating λ. We have to set the prior for the deep parameters of the model $p_0(\gamma)$ and the prior weight for the DSGE model $p_0(\lambda)$. Del Negro and Schofheide [2004] take a set of discrete grid values $\lambda \in \bar{\Lambda}$ and chose the one $\hat{\lambda}$ that maximizes the marginal data density $\hat{\lambda} = \arg\max_{\lambda \in \Lambda} p_\lambda(F)$. So this approach can be considered as a model choice problem where each model is defined by the respective value of λ^m where $m \in \{1, ..., M\}$ defines the different values and can be regarded as different models. Here opposed to DS (2004) I prefer to treat λ as a random parameter to be estimated. Hence this approach is more Bayesian in nature, it averages over the sampled values of λ and is at least computationally more efficient. I follow the arguments of Adjemaine et. al. [2007,2008] who have implemented this approach for the DSGE-VAR case in DYNARE version 4.

[8] For the singularity argument see Sargent [1989]. In case of stochastic singularity and an improper prior, i.e. $\lambda < \frac{K(P+1)}{T}$ estimation via OLS would not be possible, as in this case we would have not more observations than parameters to estimate.

Marginal posterior density of the DSGE model The post-sample information about the DSGE model parameters is summarized in the marginal posterior density of γ

$$p(\gamma \mid F^T) = \int p(\gamma \mid B, \Sigma) p(B, \Sigma \mid F^T) d(B, \Sigma) \tag{4.3.18}$$

Under the improper prior $\lambda = 0$ the DFM and the DSGE model are independent *a priori*. From the independence of γ follows that there is no updating and hence no learning from the structural parameters of the DSGE.

Marginal Data density Factors. Start by the quasi-likelihood function

$$p^*(F^T \mid \gamma) \propto |\Sigma^*(\gamma)|^{-\frac{T}{2}} \exp\{-\frac{1}{2} tr[\Sigma^{*-1}(\gamma)(F - ZB^*(\gamma))'(F - ZB^*(\gamma))]\} \tag{4.3.19}$$

To determine the marginal data density

$$p_\lambda(F^T) = \int p_\lambda(\gamma \mid F^T) p(\gamma) d\gamma \tag{4.3.20}$$

$$p_\lambda(F^T) = \int p(F^T \mid \mathbf{B}, \Sigma) p_\lambda(\mathbf{B}, \Sigma \mid \gamma) p(\gamma) d(\mathbf{B}, \Sigma, \gamma). \tag{4.3.21}$$

4.4 Identification of Shocks

One of the key objectives of this paper is to characterize the Euro area transmission mechanism of shocks to provide insights how the member countries react to Euro area wide aggregate shocks and about potential heterogeneity across the countries. This has e.g. important implications for the single monetary policy. This task is facilitated by calculating and comparing impulse response functions to different shocks across countries. The identifying assumptions are key to the results. To exemplify the identification problem I will first explain the generic identification problem related to picking a unique rotation matrix that maps the reduced form residual into identified structural shocks. Subsequently I will explain the different identification schemes I employ.

There are essentially three different identification approaches that I employ and analyze. First a pure sign restriction approach where the sign of some DSGE impulse responses are imposed on the Impulse response functions of the estimated dynamic factor model for a specified period of time and the shocks of interest. Second I employ the DSGE model rotation approach. Conditional on a draw of the deep parameters a unique rotation matrix $\Omega^{DSGE}(\gamma)$ exists that maps the Cholesky decomposition of the factors state residuals (reduced form innovations) variance-covariance into the matrix of DSGE impulse response functions on impact. Third the sign restriction approach within the DSGE-DFM approach. The distinction between the pure sign restriction approach is that a prior for the estimation of the reduced form parameters based on the theoretical second moments of the DSGE model are included. So it follows the same approach as in the DSGE-DFM approach except that for the rotation matrix the sign restriction derived from the DSGE estimation are invoked rather then the implied rotation of the DSGE model. The advantage here is that the we don't have to identify all the shocks, the model does not need to be fully stochastically defined and the ordering of the variables is not relevant.

Calculating Impulse Response Functions. In order to get the dynamic responses of the single indicators we have to impose the identifying restrictions of interest, calculate the dynamic responses of the factors to the respective structural shocks of interest. For the impulse response functions of the single variables we combine the factor responses with the respective row of the factor loading matrix which relates factors to indicators. The impulse response function of the factor k to shock s at horizon h is given by

$$r_{s,k}(0) = A_s \nu_t \tag{4.4.1}$$

$$r_{s,k}(h) = B^h r_{s,k}(0) \quad for \quad s = 1, \ldots, S. \tag{4.4.2}$$

In order to get the impulse response function of the respective indicators $r_{s,k}(h)$ has to be combined with the respective factor loadings. The resulting impulse response function of indicator n to shock s at horizon h results in

$$r_s^n(h) = \Lambda_n' r_s(h) \quad for \quad s = 1, \ldots, S \tag{4.4.3}$$

where $r_s(h)$ denotes the responses of the factors F to the impulse matrix A. The relation between the reduced form innovations u_t and the structural shocks ν_t is given by

$$u_t = A_0^{chol} \Omega \nu_t, \quad \nu_t \sim \mathcal{N}(0, I_K) \tag{4.4.4}$$

where A_0^{chol} is the Cholesky decomposition of the covariance matrix of the factor residuals and Ω is an orthonormal matrix. Identification requires a unique choice of Ω as the data is salient about the choice of Ω since the following relation holds

$$A_0^{chol} \Omega \Omega' A_0^{chol} = A_0^{chol} A_0^{chol\prime}. \tag{4.4.5}$$

The likelihood is invariant to the choice of Ω. Macroeconomist often have an *ex ante* justification to produce *ex post* reasonable results. There is a rich literature on different identification strategies in the VAR framework. For the dynamic factor model case Bernanke Boivin and Eliasz [2005] propose a recursive Cholesky identification scheme for contractionary monetary policy shocks. However there are problems and puzzles associated with this approach discussed in Amir Ahmadi and Uhlig [2].

4.4.1 DSGE Model Based Sign Restriction.

For the sign restriction approach nested in the generic description we need the following restriction. We need to find the impulse matrix $\Omega_j^{SR}(\gamma)$ which is a submatrix of the orthonormal rotation matrix $\Omega(\gamma)$. Note that the dependency of γ shows the connections to the estimated DSGE model such that results have to be in line with the derived restrictions from the estimated impulse response function of the DSGE model.

$$\Sigma_{tr}^{SR} = chol(\Sigma) \tag{4.4.6}$$

$$\Omega_j^{SR} \supseteq \Omega^{SR}(\gamma) \tag{4.4.7}$$

$$A^{SR}(\gamma) = \Sigma_{tr}^{SR} \Omega_j^{SR}(\gamma) \tag{4.4.8}$$

Identification of structural shocks through imposing sign restrictions is based on the assumptions about the sign of the impulse response functions for a specified period of key macroeconomic variables. Such restrictions should represent "conventional wisdom" derived from economic theory that most researchers can agree on. The sign restriction approach seems reasonable and promising, especially in the context of DFM in which identification restrictions can be and tightened along the lines of conventional wisdom as there are far more relevant indicators of interest.

Definition 1 *The matrix* ${A^{SR}}_j(\gamma) \in \Re^{(K)}$ *is called an* ***i****mpulse matrix, iff there is some matrix* $A^{SR}(\gamma)$*, so that* $A^{SR}(\gamma)A^{SR}(\gamma)' = \Sigma$ *and so that* ${A^{SR}}_j(\gamma)$ *is a subset of* $A^{SR}(\gamma)$*.*

4.4.2 DSGE-DFM identification

Del Negro and Schorfheide [12] according to their DSGE-VAR approach rely on the fact that the DSGE model is identified as for each set of deep parameters γ_M a unique Matrix $\Omega(\gamma_M)$ exists that maps the Cholesky decomposition of the variance-covariance matrix of forecast errors into the matrix of DSGE impulse responses on impact. As described before the relation between the reduced and structural DFM is connected via the relation between the matrix of contemporaneous relations A_0 and Σ. The DSGE implies a matrix $A(\gamma_M)$ that connects the DSGE implied reduced form DFM to its implied structural version. DS apply a QR-decomposition to $A(\gamma_M)$ denoting

$$A(\gamma_M) = A_0^* \Omega(\gamma_M) \tag{4.4.9}$$

where A_0^* is triangular and $\Omega(\gamma_M)$is orthonormal. Next they set

$$A_0 = A_0^{chol}\Omega(\gamma_M), \quad A_0^{chol*} = \mathrm{chol}(\Sigma). \tag{4.4.10}$$

Hence $(B, \Sigma, \Omega^*(\gamma_M), \Lambda, R, \Psi)$ defines an identified structural DFM.

4.5 The Theoretical Model for the Euro area

To apply the proposed procedure I will employ a state-of-the-art DSGE model based on macroeconomic foundations taken from Smets and Wouters. It is a natural choice due to the fact that it fits well to the aggregate Euro area data. Their model has gained particular interest due to its success in fitting actual data both in the US and in the Euro area. In addition it performs particularly well in terms of out-of-sample predictions compared standard VAR and Bayesian VAR models. The model involves optimizing households that consume goods and services, supply specialized labor on a monopolistically competitive labor market, rent capital services to firms and decide how much capital to accumulate. Firms choose the desired level of labor and capital inputs, and supply differentiated products on a monopolistically competitive goods market. Prices and wages are re-optimized at random intervals as in the Calvo model. Otherwise prices are partially indexed to past inflation rates.

I estimate the model including seven key macroeconomic variables for each country: real GDP, (real) consumption, (real) investment, inflation, real wages, employment and the nominal interest rate. The selected countries are Austria, Belgium, Finland, France, Germany, Italy, Ireland, Netherlands, Portugal, and Spain. In addition the Euro area aggregate data that are also used by Smets and Wouters to estimate their model are included to serve as a normalization for the factor identification.

For the paper at hand I estimate a slightly modified version of the DSGE model presented in Smets and Wouters. First I reduce the number of shocks in the original model from 10 to 7 to match the number of variables per country considered. Second, I change some inconsistencies that are present either in the working paper version, the published version and even in the code. These changes are partly documented in

Onatski and Williams [2004]. The changes I rely on are based on Uhlig's lecture notes and his implementation of the calibrated version of the model[9]. Before estimation, the following changes have been made to the model described in the next subsection.

4.5.1 Modifications to the Model

Reduction of shocks. First we reduce the original number of ten shocks down to seven in order to have as many shocks as observation in the system. The omitted shocks are the inflation objective shock $\bar{\pi}_t$, the government spending shock ϵ_t^G and the equity premium shock η_t^Q.

List of changes. The changes follow some difficulties and inconsistencies that arise from the original formulation of the paper and its working paper version. The labor supply shock in the wage equation enters with a positive sign instead of a negative sign to ensure that an increase in the labor supply decreases wages which is in line with conventional wisdom. In addition the time $(t+1)$ preference shock ϵ_{t+1}^b and the (negative) investment shock ϵ_{t+1}^I in equation (28) and (29) respectively in the original paper are taken out.

4.5.2 Log-linearized Model

For the paper to be self contained the log-linearized equations are listed here. A complete and thorough description can be found in the original paper.

The capital accumulation equation is given by

$$\hat{K}_t = (1-\tau)\hat{K}_{t-1} + \tau\hat{I}_{t-1} \tag{4.5.1}$$

The labor demand equation is given by

$$\hat{L}_t = -\hat{w}_t + (1+\psi)\hat{r}_t^k + \hat{K}_{t-1} \tag{4.5.2}$$

The goods market equilibrium condition is given by

$$\hat{Y}_t = (1-\tau k_y - g_y)\hat{C}_t + \tau k_y \hat{I}_t + \epsilon_t^G \tag{4.5.3}$$

The production function is

$$\hat{Y}_t = \phi\epsilon_t^a + \phi a\hat{K}_{t-1} + \phi a\psi\hat{r}_t^k + \phi(1-\alpha)\hat{L}_t \tag{4.5.4}$$

The monetary policy reaction function is given by

$$\begin{aligned}\hat{R}_t = \quad & \rho\hat{R}_{t-1} + (1-\rho)\left[\bar{\pi}_t + r_\pi(\hat{\pi}_{t-1} - \bar{\pi}_t) + r_Y(\hat{Y}_t - \hat{Y}_t^P)\right] \\ & + r_{\Delta\pi}(\hat{\pi}_t - \hat{\pi}_{t-1}) + r_{\Delta Y}\left[\hat{Y}_t - \hat{Y}_t^P - (\hat{Y}_{t-1} - \hat{Y}_{t-1}^P)\right] + \eta_t^R \end{aligned} \tag{4.5.5}$$

The consumption equation is given by

$$\hat{C}_t = \frac{h}{1+h}\hat{C}_{t-1} + \frac{1}{1+h}E_t\hat{C}_{t+1} - \frac{1-h}{(1+h)\sigma_c}(\hat{R}_t - E_t\hat{\pi}_{t+1}) + \frac{1-h}{(1+h)\sigma_c}\hat{\epsilon}_t^b \tag{4.5.6}$$

[9] For the manual, lecture notes and the model implementation in TOOLKIT see http://www2.wiwi.hu-berlin.de/institute/wpol/html/toolkit/MacAppSoft/MacroApplSoftware2.1.html

The investment equation is given by

$$\hat{I}_t = \frac{1}{1+\beta}\hat{I}_{t-1} + \frac{\beta}{1+\beta}E_t\hat{I}_{t+1} + \frac{\varphi}{1+\beta}\hat{Q}_t + \hat{\epsilon}_t^I \tag{4.5.7}$$

The Q equation is given by

$$\hat{Q}_t = -(\hat{R}_t - E_t\hat{\pi}_{t+1}) + \frac{1-\tau}{1-\tau-\bar{r}^k}E_t\hat{Q}_{t+1} + \frac{\bar{r}^k}{1-\tau-\bar{r}^k}E_t\hat{r}_t^k + \eta^Q \tag{4.5.8}$$

The inflation equation is given by

$$\begin{aligned}\hat{\pi}_t = \quad & \frac{\beta}{1+\beta\gamma_p}E_t\hat{\pi}_{t+1} + \frac{\gamma_p}{1+\beta\gamma_p}\hat{\pi}_{t-1} \\ & + \frac{1}{1+\beta\gamma_p}\frac{(1-\beta\xi_p)(1-\xi_p)}{\xi_p}[\alpha\hat{r}_t^k + (1-\alpha)\hat{\omega}_t - \hat{\epsilon}_t^a] + \eta_t^p\end{aligned} \tag{4.5.9}$$

The wage equation follows:

$$\begin{aligned}\hat{w}_t = \quad & \frac{\beta}{1+\beta}E_t\hat{w}_{t+1} + \frac{1}{1+\beta}\hat{w}_{t-1} + \frac{\beta}{1+\beta}E_t\hat{\pi}_{t+1} - \frac{1+\beta\gamma_w}{1+\beta}\hat{\pi}_t + \frac{\gamma_w}{1+\beta}\hat{\pi}_{t-1} \\ & - \frac{1}{1+\beta}\frac{(1-\beta\xi_w)(1-\xi_w)}{(1+\frac{(1+\lambda_w)\sigma_L}{\lambda_w})\xi_w}[\hat{w}_t - \sigma_L\hat{L}_t - \frac{\sigma_c}{1-h}(\hat{C}_t - h\hat{C}_{t-1}) + \hat{\epsilon}_t^L] + \eta_t^w\end{aligned} \tag{4.5.10}$$

Summary of the model. The modified model has seven orthogonal shocks, including productivity shock, preference shock, labor supply shock, investment shock, monetary policy shock, price mark-up shock and a wage mark-up shock. The vector of exogenous shocks is given by:

$$\epsilon_t = (\epsilon_t^a, \epsilon_t^b, \epsilon_t^L, \epsilon_t^I, \eta_t^R, \eta_t^p, \eta_t^w) \tag{4.5.11}$$

The vector of endogenous variables is given by

$$z_t = (Y_t^*, C_t^*, I_t^*, K_t^*, Q_t^*, r_t^{k*}, \pi_t^*, w_t^*, i_t^*, Y_t, C_t, I_t, K_t, Q_t, r_t^k, \pi_t, w_t, i_t) \tag{4.5.12}$$

The vector of predetermined endogenous or lagged exogenous variables is given by

$$Z_t = (Y_{t-1}, C_{t-1}, I_{t-1}, K_{t-1}, Q_{t-1}, \pi_{t-1}, w_{t-1}, i_{t-1}). \tag{4.5.13}$$

4.6 Empirical Results

4.6.1 Data

I estimate the model using quarterly data on GDP, consumption, investment, real wage, employment, inflation, and short-term interest rate. I consider Euro area data and country data from Austria, Belgium, Finland, France, Germany, Ireland, Italy, Netherlands, Portugal and Spain. The sample starts in Q1 1985 and ends in Q4 2007. Note that the cross-sectional dimension of the panel is 76[10], and the time-series dimension of the panel is 92.The sources for the data are DataStream and EcoWin. We collected data on GDP, consumption, investment, nominal wage, employment, GDP deflator,

[10]Data on wages for Belgium is missing.

and short-term interest rate.[11] The data is transformed as follows. Whenever a time series was not seasonally adjusted, we deseasonalized the time series using the X11 command in RATS. We took logs of GDP, consumption, investment, and employment. We computed the log of real wage as the log of nominal wage minus the log of the GDP deflator. We computed inflation on a year-on-year basis as the difference of the log of GDP deflator from the quarter of the previous year. We divided interest rates by 4 to express interest rates in quarterly terms. We detrended each variable by its own linear trend, except that we detrended interest rates by the same trends as inflation rates. We multiplied each variable by 100, except for interest rates. This yielded a dataset with each variable expressed in percentage point deviation from trend, with mean zero, and with unnormalized variance. Note that Smets and Wouters use the same euro area variables as we do, and transform the variables in the same way as we do[12]. Smets and Wouters use only euro area data. The sample in Smets and Wouters starts in Q1 1970 and ends in Q4 1999. Some data from Germany that we collected display a discrete change in the first quarter of 1991. When detrending each time series from Germany, we allowed for two separate trends: one trend from Q1 1985 to Q4 1990, and another trend from Q1 1991 to Q4 2007.

4.6.2 Data Analysis

Is there a factor structure? To judge whether the tool of dynamic factor analysis is appropriate it is necessary to ensure that there is a factor structure in the data. For that I report results of principal component analysis for each of the country data and the full data set. In particular based on a pareto plot that shows the explained variation of the data by each principal component and it's cumulative sum. For the data to have a factor structure few principal components should explain a key part of the variation in the data. This is the case as can be seen in figure (4.1).

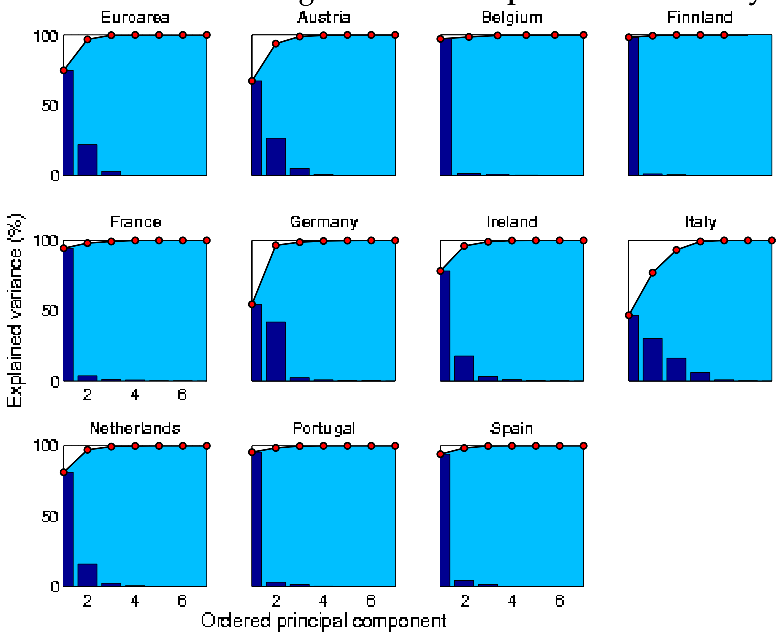
4.6.3 Model Specification

The number of factors K is a-priori determined to be 7. This choice is due the fact that the DSGE model considered has seven shocks. The DSGE-DFM estimation and identification approach requires to have as many factors as shocks which is a disadvantage towards the sign restriction approach. However to facilitate the comparability between the approaches and to explore the proposed approach I will fix the number of factors. It turns out that the improvement in the fit by increasing the number of factor is negligible. Hence K could be reduced with a comparable fit therefore fixing this parameter turns out not to be restrictive for the analysis. Results are reported for $P = 4, Q = 2$ lags. The number of Gibbs iterations is 100.000, with a burn in of 50.000 and a thinning parameter of 10 resulting in 5000 kept draws. Convergence of the sampler is monitored until the target distribution is reached.

[11] GDP, consumption, and investment were reported in real terms. Employment was reported as the number of persons.

[12] Except for inflation they calculate a quarter-on-quarter inflation series. It turns out that the version with year-on-year inflation delivers a much better fit.

Figure 4.1: **Pareto plot for the country data.**



4.6.4 Model Fit

To assess how well the model and the factors fit the data I calculate the shares of the variance in the single data series explained by the factors. Most variables are rather well explained. However the results depend on two restricting model specification. First results depend on the restrictions imposed on the factor loading matrix to uniquely identify the factors against rotational indeterminacy. The block diagonally assumption is required to have a structural interpretation of the factor and to facilitate the Factor-DSGE estimation part. The strong restriction results in slightly worse fit but still the overall explained variation is high regarding most of the real indicators which is not in conflict with the conclusion of a high degree of comovement in the Euro area data. However the explained variation in the inflation series is much worse hence calculating the impulse response functions for those variables is less meaningful. Secondly considering a year-on-year inflation series rather than quarter-on -quarter affects the fit measured by the explained variation of the inflation series by much. For the reduced form DFM estimation with the factor identification imposing the upper $K \times K$ block set to identity I deliver the best results in terms of model fit measured by the R^2 statistic. The variation of around $2/3$ of the data series are explained by more than (90%) as can be seen in table (4.4). The series explained the least are the inflation indicators. The results suggest a factor structure that is very well explained by the factors.

4.6.5 Model Comparison

In order to evaluate how well the estimated models fit the data I report the calculation of the marginal likelihood of the models based on the Laplace approximation and alternatively on the modified harmonic mean estimator proposed by Geweke [1998]. In

Table 4.4: **Macro Variables and share of variance explained by estimated factors**

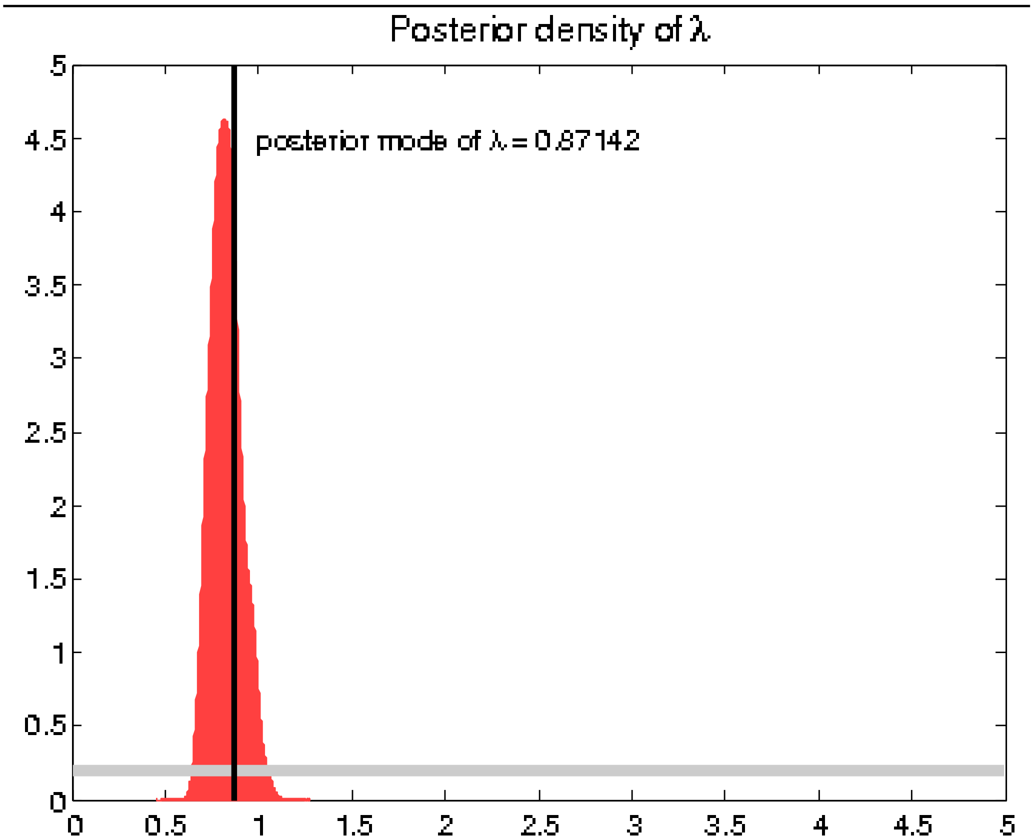
Description	R^2	Description	R^2
ES Employment	1.00	IE Real consumption	0.94
FI Real consumption	0.99	AT Real GDP	0.94
FI Employment	0.99	BE Real consumption	0.93
EU Employment	0.99	PT Real consumption	0.93
FI Real GDP	0.99	IT Real consumption	0.93
IT Employment	0.99	IE Real GDP	0.92
FR Real gross investment	0.98	PT Employment	0.92
FI Real gross investment	0.98	FR Employment	0.92
AT Short-term interest rate	0.98	IT Real gross investment	0.90
EU Short-term interest rate	0.98	BE Real GDP	0.89
NL Short-term interest rate	0.98	NL Real gross investment	0.88
BE Employment	0.98	BE Real gross investment	0.88
ES Real consumption	0.98	DE Employment	0.86
PT Real GDP	0.98	NL Real wage	0.86
FR Short-term interest rate	0.98	IT Real GDP	0.86
ES Short-term interest rate	0.97	IE Real wage	0.86
NL Real GDP	0.97	AT Real gross investment	0.85
NL Employment	0.97	FI Real wage	0.84
ES Real wage	0.97	IE Real gross investment	0.84
ES Real GDP	0.97	DE Real gross investment	0.84
FI Short-term interest rate	0.97	EU Inflation	0.83
DE Short-term interest rate	0.97	AT Real wage	0.82
PT Short-term interest rate	0.97	FR Real wage	0.81
ES Real gross investment	0.97	IE Short-term interest rate	0.81
IT Short-term interest rate	0.97	AT Inflation	0.79
FR Real GDP	0.97	DE Real consumption	0.77
BE Short-term interest rate	0.96	DE Inflation	0.76
EU Real gross investment	0.96	FR Inflation	0.75
IE Employment	0.96	PT Inflation	0.75
PT Real gross investment	0.96	ES Inflation	0.72
FR Real consumption	0.96	DE Real GDP	0.71
NL Real consumption	0.96	DE Real wage	0.68
AT Employment	0.95	NL Inflation	0.62
PT Real wage	0.95	IT Real wage	0.56
EU Real consumption	0.95	FI Inflation	0.51
AT Real consumption	0.95	IT Inflation	0.44
EU Real wage	0.94	BE Inflation	0.44
EU Real GDP	0.94	IE Inflation	0.34

order to compare the models I furthermore report the Bayes factor. It emerges immediately that the combined estimation of the DFM with the DSGE model fits better DSGE model the pure DSGE estimation approach. This hold for both approaches to calculate the marginal data density. However the model version where the DSGE model is estimated with the Euro are aggregates as opposed to the extracted factors for the respective economic concepts derive from all the data of all countries fits better. Hence the best model is the DSGE-DFM approach based on the estimated DSGE model that is based on the Euro area aggregates. How to employ the modified harmonic mean estimator and alternative approaches proposed by Sims, Waggoner and Zha [2008] and a more general bridge sampling approach of Meng and Wong can be found in the appendix.

4.6.6 DSGE prior weight

How much do we learn from the DSGE model and what is it's optimal value? The key metric for that is the DSGE prior weight explained before and denoted by λ. It distribution can be found in figure (4.2). The posterior mode value is $\lambda_{mode} = .87$ which is close but slightly less then the applications of Del Negro and Schorfheide [12] who find for their application to the post-world war II US data optimal values of $\lambda \in \{.75, 1, 125\}$. Note that I assume a rather uninformative uniform prior on the interval $[0, 5]$. Even much wider intervals like $[0, 100]$ resulted in very similar posterior distribution for the weight λ which is narrowly peaked around the posterior mode. Therefore I stick to the lower range interval as the choice of the optimal λ and it's posterior distribution is not affected. This way the integration on low likelihood parameters is avoided and hence the simulation of the marginal data density is not artificially deteriorated[13].

Figure 4.2: **Posterior distribution of Model weight:** λ



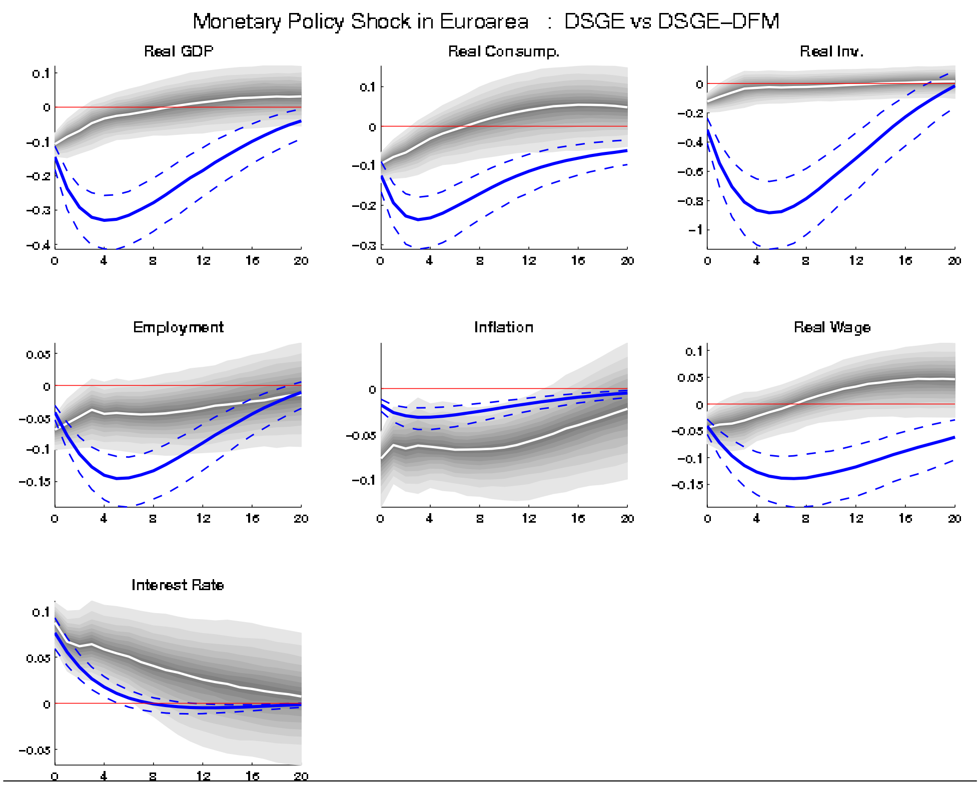
[13] Adjemian et. al [2008].

4.6.7 Impulse Response function Analysis: Monetary Policy Shock

4.6.7.1 Comparison: DSGE-DFM vs. DSGE

In order to detect the difference in the monetary transmission mechanism implied by the pure DSGE model and the DSGE-DFM I plot the respective calculated impulse response functions against each other in figure (4.3). The reaction of the real indicators such as real GDP, real wages and in particular real investment is significantly stronger than implied by the DSGE-DFM. In contrast the impulse response function of inflation is less strong. Overall one can observe that the DSGE impulse response functions have a hump-shaped response between 4-6 quarters. An exception is the interest rate response which is strongest on impact and afterwards dies out to the pre-shock level after 8 quarters. The persistence in the DSFE-DFM is less for all variables except for interest rates where the reaction is stronger with respect to the scale and shows a higher persistence.

Figure 4.3: **IRF of DSGE-DFM-Indicator**



4.6.7.2 Identification of DSGE-DFM via DSGE rotation with block-diagonal Λ

As noted before the DSGE rotation based identification of the DSGE-DFM relies on the block-diagonally restrictions imposed on the factor loading matrix to uniquely identify the factors against rotational indeterminacy. Hence by construction there is homogeneity in the shape of the impulse response functions as the are multiple scalars of the factor impulse response functions. However differences in the scale can be analyzed which gives insights to the exposure of the single member countries to aggregate shocks. In general I find that this approach is associated with a rather high degree of uncertainty compared to the DSGE-DFM case identified with sign restrictions. The results reported suggest that the impulse response functions of most countries real GDP to a contractionary euro area aggregate monetary policy shock is close to the response of the Euro area GDP factor as regards the scale. Austria and Portugal show a weaker response indicating a lower exposure to the shock. In terms of frequentist language the results do not show a significant price puzzle however there is some probability mass on the positive support. This might be unappealing and we might want to restrict or penalize a positive price response to a contractionary monetary policy. This is feasible and I employ it in the sign restriction (both in the DSGE-DFM and the DFM approach) version in the next two subsections. The strongest discrepancy regarding the scale of the impulse response functions can be found in the real wages and employment in Finland, Portugal and Spain. The reaction is up to ten mimes stronger than in Germany or France. Overall the conclusion of DSGE-DFM approach with the DSGE rotation is that there is by and large homogeneity in the monetary transmission mechanism though there are some differences in the exposure of some countries' labor market to a contractionary monetary policy shock.

4.6.7.3 Identification of DSGE-DFM via Sign Restriction with block diagonal Λ

The restrictions imposed are derived from the estimated DSGE model and results in the sign restrictions reported in table (4.5). Following Uhlig [2005] I impose the restrictions for the contemporaneous period and up to two quarters. Compared to the results of the previous section employing the DSGE rotation to facilitate identification the results of the sign restriction approach within the DSGE-DFM are associated with less uncertainty and the responses show a higher degree of homogeneity regarding the respective scales of the impulse response functions following a contractionary monetary policy shock. Most countries real GDP responses are close to the Euro area GDP factor impulse response function. Slightly weaker responses can be found for Austria and stronger responses in Portugal. As discussed in the previous section the prize puzzle is ruled out by construction without any probability mass on the positive support. Furthermore the single countries inflation responses are all similar to the Euro area responses. The price puzzle of the previous section is ruled out by construction. However the scale if the responses are rather similar here. The same picture regarding the strongest discrepancy in the scale of the responses emerges for the reaction of employment and real wages across countries. In Finland, Portugal and Spain the reaction is up to 5 times stronger compared to Germany or France. The impulse responses to the interest rates look more reasonable compared to the previous case in that they have a clear positive impact in the initial quarters following the shock. The uncertainty in the previous case is very high.

Table 4.5: **Model-based robust sign restrictions**

	η^w	ε^a	η^p	η^R	ε^L	ε^I	ε^b
Real GDP	≤ 0	$0 \geq$	≤ 0	≤ 0	$0 \geq$	$0 \geq$	$0 \geq$
Real consumption	≤ 0	$0 \geq$	≤ 0	≤ 0	$0 \geq$	≤ 0	$0 \geq$
Real investment	≤ 0	$0 \geq$	≤ 0	≤ 0	$0 \geq$	$0 \geq$	≤ 0
Employment	≤ 0	≤ 0	≤ 0	≤ 0	$0 \geq$	$0 \geq$	$0 \geq$
Inflation	$0 \geq$	≤ 0	$0 \geq$	≤ 0	X	$0 \geq$	$0 \geq$
Real wage	$0 \geq$	$0 \geq$	≤ 0	≤ 0	≤ 0	$0 \geq$	$0 \geq$
Interest rate	$0 \geq$	≤ 0	$0 \geq$	$0 \geq$	≤ 0	$0 \geq$	$0 \geq$

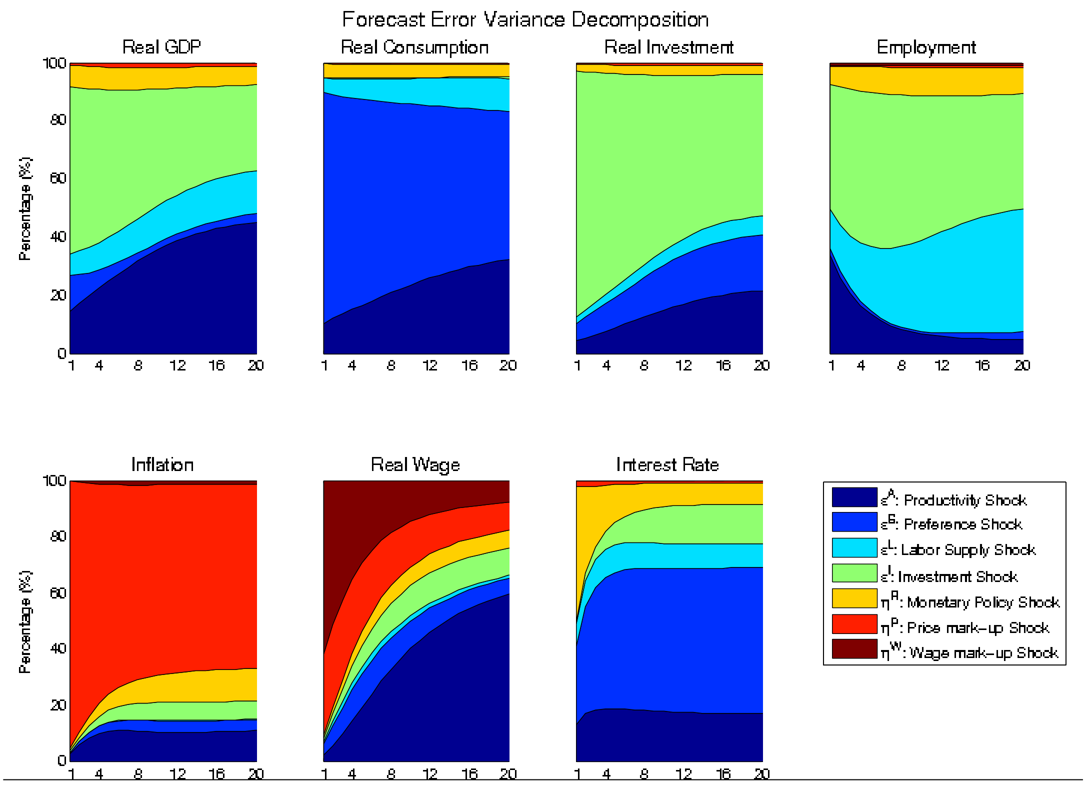
4.6.7.4 Identification of DFM via Sign Restriction

As a third alternative I employ the sign restriction scheme to identify aggregate shocks, however the DSGE model serves only as a tool to derive robust sign restrictions and NOT to estimate the parameters of the empirical model as opposed to the previous two models. This approach shows a higher degree of heterogeneity in particular at longer horizons. At short horizons most of the responses have the same sign though the shape can be somewhat different. Furthermore there a some sign puzzles. The key appeal is that this approach is less restrictive allowing for richer and more complicated dynamics in the transmission of shocks. This is due to the fact that the normalization of the loading matrix is by far less restrictive as no block-diagonally is required. The estimated DSGE model delivers the same sign restriction to be imposed both when based on the euro area aggregate data or on factors. This complexity comes at a cost. For the results to reasonable more sign restrictions might have to be imposed. From the figures we can see that there are some sign puzzles in particular regarding the reaction of prices. It should be noted that the price puzzle emerge for those variables that are rather badly represented by the data hence the figure might appear less reasonable as it actually is. It is obvious that if those indicators that are mostly driven by idiosyncratic dynamics their exposure to surprise changes driving the common factors is negligible even thought the respective impulse response functions have clear effects. Hence while interpreting the impulse response functions one should always keep in mind how much do the respective indicators depend on the common factors.

4.6.8 Forecast Error Variance Decomposition: Monetary Policy Sock

The key structural shocks explaining most of the variation in output are the productivity shock and the investment shock which explain about 75% of the variation over the entire horizon considered. The contribution of the investment shock decreases over the horizon whereas the contribution to the productivity shock increases by almost the same horizon. The monetary policy shock has a low contribution to the variation in real GDP at an average proportion of 6%-8%. Hence consistent with much of the VAR literature for the US data monetary policy shock play a minor role for generating business cycles. This holds true for most of the variables considered except for interest rates where the impact contribution is around 50% decreasing rather quickly after 8 quarters to around 10%. Labor supply shocks turn out to have an explanatory power in the variation in employment only, price mark up shocks are important only for the variation in inflation and wage mark-up shocks only for real wages in the short-run. With increasing horizon the productivity shock gains a more crucial importance to explain the variation in real wages.

Figure 4.4: **Forecast error variance decomposition of DSGE-DFM**



4.7 Conclusion

In this paper I seek to answer two questions with the help of a new proposed methodology combining the strength of recent advances in empirical macroeconomics. First I ask what is the degree of comovement in the macroeconomic dynamics of the Euro area and it's member countries. I find a high degree of comovement among the countries in particular regarding the real side of the economy. Depending on the transformation of the inflation series the comovement can be less strong. Considering year-on year inflation the comovement is higher compared to quarter-on-quarter inflation.

The second question I pose is whether there is heterogeneity in the transmission mechanism of macroeconomic shocks. With an emphasis on contractionary monetary policy shocks within the Euro area I find by and large a similar transmission mechanism across the constituent countries though there are some differences regarding the labor market in the countries. Real wages and employment in Germany and France are less affected whereas the exposure of these indicators in Finland, Portugal and Spain is between 5-10 stronger.

Monetary policy shocks play a minor role for business cycle fluctuations resulting in a constant average of about 8% over the entire horizon. The key drivers are productivity and investment shocks that contribute over the entire horizon of around 75% with an increasing contribution of productivity shocks and a decreasing contribution of investment shocks along the horizon.

The methodology I propose in this paper draws on the recent advances in empirical macroeconomics combining Bayesian DFM with state of the art DSGE estimation , the hybrid combination thereof (Del Negro and Schorfheide [12] and Sims [2008]) and the identification of macroeconomic shocks. The resulting DSGE-DFM is promising as it delivers structural factors, allows for theory based identification formulated by fully fledged and well fitting DSGE models. Based on formal model comparison the DSGE-DFM always outperforms the DSGE model independently of the comparison based on indicators or factors representing the economic concepts of the DSGE resulting in higher marginal likelihood. Hence the restrictions implied by the DSGE model are supported by the data improving the fit. This is confirmed by the positive value of the data driven posterior value (and it posterior distribution) of the DSGE model weight around 0.87. Hence I conclude that it is helpful to combine DFM estimation with well fitting DSGE models.

Regarding the identification of shocks the imposed DSGE rotation is appealing as the structural DSGE model is exactly identified. However depending on the application at hand employing the sign restriction approach might be of particular interest as it is more flexible in three dimensions. First it can be employed even when there is no well fitting DSGE model. It does not depend on the joint estimation of the DSGE model. If it fits bad hence the DSGE rotation might not be of particular appeal and reliable to impose. Furthermore in large dimensional models there are many more indicators to be restricted if the researcher has some sounds a-priori reasoning (see Amir Ahmadi and Uhlig [2]). And finally the third enticing promise is that not all shocks have to be identified. If the researcher is interested in only one or a subset of shocks she can do so without depending on a fully stochastically defined DSGE model. The empirical results based on the sign restriction in the DSGE-DFM shows a lower degree of uncertainty compared to the identification based on the DSGE rotation.

Appendix C

Appendix C.1 Tables

Table 4.6: **Forecast Error Variance Decomposition (FEVD).**

Horizon	Shock	Y	C	I	E	π	W	R
h=0	Productivity Shock	0.15	0.10	0.04	0.34	0.03	0.02	0.13
	Preference Shock	0.12	0.79	0.06	0.02	0.00	0.04	0.28
	Labor Supply Shock	0.07	0.05	0.02	0.13	0.00	0.01	0.08
	Investment Shock	0.58	0.00	0.85	0.43	0.01	0.01	0.01
	Monetary Policy Shock	0.07	0.05	0.02	0.06	0.01	0.01	0.48
	Price mark-up Shock	0.01	0.00	0.00	0.00	0.95	0.29	0.02
	Wage mark-up Shock	0.00	0.00	0.00	0.01	0.00	0.62	0.00
h=1	Productivity Shock	0.17	0.12	0.05	0.26	0.06	0.05	0.17
	Preference Shock	0.10	0.77	0.07	0.02	0.01	0.07	0.38
	Labor Supply Shock	0.08	0.06	0.02	0.16	0.00	0.02	0.09
	Investment Shock	0.56	0.00	0.82	0.48	0.01	0.03	0.03
	Monetary Policy Shock	0.08	0.05	0.03	0.07	0.02	0.02	0.30
	Price mark-up Shock	0.01	0.00	0.01	0.00	0.89	0.29	0.02
	Wage mark-up Shock	0.00	0.00	0.00	0.01	0.01	0.51	0.00
h=4	Productivity Shock	0.25	0.17	0.09	0.14	0.11	0.19	0.19
	Preference Shock	0.05	0.71	0.10	0.01	0.03	0.12	0.49
	Labor Supply Shock	0.10	0.07	0.03	0.22	0.00	0.02	0.10
	Investment Shock	0.51	0.00	0.73	0.53	0.04	0.07	0.08
	Monetary Policy Shock	0.08	0.05	0.03	0.09	0.06	0.05	0.13
	Price mark-up Shock	0.02	0.01	0.01	0.01	0.75	0.24	0.01
	Wage mark-up Shock	0.00	0.00	0.00	0.01	0.02	0.30	0.00
h=8	Productivity Shock	0.34	0.22	0.14	0.07	0.11	0.36	0.18
	Preference Shock	0.03	0.64	0.14	0.01	0.04	0.11	0.51
	Labor Supply Shock	0.12	0.08	0.05	0.30	0.00	0.02	0.09
	Investment Shock	0.42	0.00	0.63	0.51	0.06	0.10	0.12
	Monetary Policy Shock	0.08	0.04	0.03	0.10	0.09	0.07	0.09
	Price mark-up Shock	0.02	0.01	0.01	0.01	0.68	0.17	0.01
	Wage mark-up Shock	0.00	0.00	0.00	0.01	0.02	0.17	0.00
h=16	Productivity Shock	0.44	0.30	0.21	0.05	0.11	0.56	0.17
	Preference Shock	0.03	0.53	0.19	0.02	0.04	0.06	0.52
	Labor Supply Shock	0.14	0.11	0.06	0.41	0.00	0.01	0.09
	Investment Shock	0.31	0.00	0.50	0.41	0.06	0.10	0.14
	Monetary Policy Shock	0.07	0.04	0.03	0.09	0.11	0.06	0.08
	Price mark-up Shock	0.01	0.01	0.01	0.01	0.66	0.11	0.01
	Wage mark-up Shock	0.00	0.00	0.00	0.01	0.02	0.09	0.00

Table 4.7: **Forecast Error Variance Decomposition (FEVD).**

Horizon	**Shock**	**Y**	**C**	**I**	**E**	π	**W**	**R**
h=20	Productivity Shock	0.45	0.33	0.22	0.05	0.11	0.60	0.17
	Preference Shock	0.03	0.50	0.19	0.03	0.04	0.06	0.52
	Labor Supply Shock	0.15	0.11	0.07	0.43	0.00	0.01	0.09
	Investment Shock	0.29	0.01	0.48	0.39	0.06	0.10	0.14
	Monetary Policy Shock	0.06	0.04	0.03	0.09	0.12	0.06	0.08
	Price mark-up Shock	0.01	0.01	0.01	0.01	0.65	0.10	0.01
	Wage mark-up Shock	0.00	0.00	0.00	0.01	0.02	0.08	0.00

Appendix C.2 Gibbs Sampling

The idea of Gibbs sampling goes back to the work of Geman and Geman [1984] and Gelfend and Smith [1990]. It aims to empirically approximate a joint posterior by blocking it into conditional distributions of known form to sample from. One produces a sequence of samples by iteratively cycling through the blocks of conditionals. Consider a parameter space can be blocked into k blocks of sub parameters or conditionals given by $\Theta = (\theta_1, \ldots, \theta_k)$. The joint density can be empirically approximated by iteratively cycling through the following steps

1. sample $\theta_1^{(g)}$ from $p(\theta_1 \mid \theta_2^{(g-1)}, \theta_3^{(g-1)}, \ldots, \theta_k^{(g-1)})$
2. sample $\theta_2^{(g)}$ from $p(\theta_2 \mid \theta_1^{(g)}, \theta_3^{(g-1)}, \ldots, \theta_k^{(g-1)})$

$\vdots$

k. sample $\theta_1^{(g)}$ from $p(\theta_k \mid \theta_1^{(g)}, \theta_2^{(g)}, \ldots, \theta_{k-1}^{(g)})$

These k steps are repeated many times until convergence has been achieved. The Gibbs sampler can be started with any set of starting values for the parameter set. An initial number of draws known as *burn-in* is discarded to avoid the influence of initial transients and the dependency on potentially bad starting values. Furthermore the sampled sequence can be reduced by a thinning parameters in order to reduce the autocorrelation of the chain. As Geman and Geman [1984] have shown the empirical distribution theoretically converges to the true invariant stationary distribution for the number of iterations large enough. Hence sampling from the conditional densities of the blocks is equivalent to sample from the joint density.

Appendix C.3 Factor DSGE estimation via MCMC

Step 1: Sample Factors

Boivin and Giannoni (2007) follow the multi-move Gibbs sampling approach by Carter and Cohn (1994), Frühwirth-Schnatter (1994) surveyed in Kim and Nelson to sample the history of the factors as the product of conditional distributions at each date t as follows:

$$p(F^T \mid \theta_M, \theta_F, X^T) = p(F_T \mid \theta_M, \theta_F, X^T) \prod_{t=1}^{T} p(F_t \mid F_{t+1}, \theta_M, \theta_F, X^T)$$

which relies on the Markov property of the F_t. Given that the state space representation in linear and Gaussian, we have

$$\begin{aligned} F_T \mid \theta_M, \theta_F, X^T &\sim \mathcal{N}(F_{T|T}, P_{T|T}) \\ F_t \mid F_{t+1}, \theta_M, \theta_F &\sim \mathcal{N}(F_{t|t,F_{t+1}}, P_{t|t,F_{t+1}}) \end{aligned}$$

where

$$\begin{aligned}
F_{T|T} =& E[F_T \mid \theta_M, \theta_F, X^T] \\
F_{T|T} =& cov(F_T \mid \theta_M, \theta_F, X^T) \\
F_{t|t,F_{t+1}} =& E[F_T \mid F_{t+1}, \theta_M, \theta_F, X^T] \\
=& F_{t|t} + P_{t|t}B'(GP_{t|t}G' + \Sigma_u)^{-1}(F_{t+1} - BF_{t|t}) \\
P_{t|t,F_{t+1}} =& cov(F_T \mid F_{t+1}, \theta_M, \theta_F, X^T) \\
=& P_{t|t} + P_{t|t}B'(BP_{t|t}B' + \Sigma_u)^{-1}BP_{t|t},
\end{aligned}$$

here the standard notation holds as in Kim and Nelson (1999) and Hamilton for the expectation conditional on the respective information set.

Step 2: $p(\Lambda, R \mid F^T, \Psi^{(g-1)}, X^T)p(\Psi \mid F^T, \Lambda^{(g)}, R^{(g)}, X^T)$

Here we have to take care of the autocorrelation in the observation equation due to the autoregressive process in the idiosyncratic process. The following equations have to be introduced:

$$\begin{aligned}
\tilde{X}_t =& X_t - \Psi X_{t-1} \\
\tilde{F}_t =& F_t - \Psi F_{t-1} \\
\tilde{X}_t =& \Lambda \tilde{F}_t + \nu_t.
\end{aligned}$$

Note that R is diagonal and hence given F^T and Ψ standard regression methods can be applied to do inference equation by equation. Here I follow closely Boivin and Giannoni (2007) by setting the prior for each equation as

$$\begin{aligned}
R_{nn} \sim& \mathcal{IG}(\delta_0, \nu_0) \\
\Lambda_n \sim& \mathcal{N}(\Lambda_0, R_{nn}^{-1} M_0^{-1})
\end{aligned}$$

where the degrees of freedom is $\nu_0 = 0.001$ the prior scale is $\delta_0 = 3$ and the prior variance in the coefficient of each observation equation is $M_0 = I_K$. The posterior distribution is given by

$$\begin{aligned}
R_{nn} \mid F^T, X^T \sim& \mathcal{IG}(\delta_T, \nu_T) \\
\Lambda_n \mid F^T, X^T \sim& \mathcal{N}(\bar{\Lambda}_n, R_{nn}^{-1} \bar{M}_0^{-1})
\end{aligned}$$

where

$$\begin{aligned}
\nu_T =& T + \nu_0 \\
\delta_T =& \delta_0 + u'u + (\Lambda_n - \hat{\Lambda}_n)'(M_0^{-1} + (Z'Z)^{-1})^{-1}(\Lambda_n - \hat{\Lambda}_n) \\
\bar{M}_n =& M_0 + (Z'Z)^{-1}
\end{aligned}$$

Give the previous draws we have $\nu_t = X_t - \Lambda F_t$ and assuming a standard normal prior

$$\Psi_{kk} \sim \mathcal{N}(0, 1)$$

the posterior is

$$\Psi_{kk} \mid F^T, \Lambda^{(g)}, R^{(g)}, X^T \sim \mathcal{N}(\bar{\Psi}_k k, \bar{N}^{-1})$$

where

$$\begin{aligned} \bar{\Psi}_{kk} &= \bar{N}^{-1} R_{kk}^{-1} \hat{v}_n' \hat{v}_n \hat{\Psi}_k k \\ \bar{N}^{-1} &= 1 + R_{kk}^{-1} \hat{v}_n' \hat{v}_n \end{aligned}$$

where $\hat{\Psi}_k k$ denotes the OLS estimate of the residuals on its lagged values.

Step 3: Sample Modelparameters

The standard way to draw from the vector of parameters is to apply a Random Walk Metropolis-Hastings algorithm as it is described in Schorfheide (2000) and An and Schorfheide (2006??). Given the previous draw $\theta_M^{(g-1)}$ a candidate draw θ_M^c is drawn from

$$\theta_M^c \sim \mathcal{N}(\theta_M^{(g-1)}, c^2 \Sigma_M)$$

where Σ_M is the inverse Hessian evaluated at the posterior mode and c is a scaling factor to ensure the appropriate acceptance rate. Based on the solution give the candidate draw the candidate draw is accepted with probability

$$\alpha(\theta_M^c, \theta_M^{g-1}) = min \left\{ 1, \frac{p(\theta_M^c \mid \Lambda, R, X^T)}{p(\theta_M^{(g-1)} \mid \Lambda, R, X^T)} \right\}$$

otherwise rejected and the draw at stage g is set to the previous $\theta_M^{(g)} = \theta_M^{(g-1)}$. We cycle through steps 1 to 3 G-times until the distribution has converged to the target distribution. In order to check the convergence of the chain a battery of convergence diagnostics are employed to ensure convergence.

Bibliography

[1] Pooyan Amir Ahmadi and Albrecht Ritschl. Monetary policy during the great depression: A bayesian favar approach. *mimeo, London School of Economics*, 2007.

[2] Pooyan Amir Ahmadi and Harald Uhlig. Measuring the effects of a shock to monetary policy: A bayesian factor augmented vector autoregression (bfavar) approach with sign identification. *mimeo*, 2010.

[3] Francesco Belviso and Fabio Milani. Structural factor-augmented vars (sfavars) and the effects of monetary policy. *The B.E. Journal of Macroeconomics*, 0(3):157–211, 2006.

[4] Ben Bernanke and Jean Boivin. Monetary policy in a data-rich environment. *Journal of Monetary Economics*, 50(3):525–546, 2003.

[5] Ben Bernanke and Kevin Carey. Nominal wage stickiness and aggregate supply in the great depression. *Quarterly Journal of Economics*, 111:853–83, 1996.

[6] Ben S. Bernanke, Jean Boivin, and Piotr Eliasz. Measuring the effects of monetary policy: A factor-augmented vector autoregressive approach. *Quarterly Journal of Economics*, 120(1):387–422, 2005.

[7] Oliver J. Blanchard and Mark Watson. Are all business cycles alike? In R. J. Gordon, editor, *The American Business Cycle*, pages 1–65. Chicago: University Press, 1986.

[8] Fabio Canova and Gianni De Nicolo. Monetary disturbances matter for business cycle fluctuations in the g-7. *Journal of Monetary Economics*, 49:1131–1159, 2002.

[9] Fabio Canova and Joaquim Pivis Pina. Monetary policy misspecification in var models. *CEPR Discussion Papers 2333*, 1:1–100, 1999.

[10] C.K. Carter and P. Kohn. On gibbs sampling for state space models. *Biometrika*, 81:541–553, 1994.

[11] Lawrence Christiano, Martin Eichenbaum, and Charles Evans. Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy*, 113:1–79, 2005.

[12] Marco Del Negro and Frank Schorfheide. Priors from general equilibrium models for VARs. *International Economic Review*, 45(2):643–673, 2004.

[13] Piotr Eliasz. Likelihood-based inference in large dynamic factor models using gibbs sampling. *Princeton University, unpublished Working Paper*, 2005.

[14] John Faust. The robustness of identified var conclusions about money. *International Finance Discussion Papers 610*, 1:1–100, 1998.

[15] Milton Friedman and Anna Jacobson Schwartz. *A Monetary History of the United States, 1867-1960*. Princeton University Press, 1963.

[16] Sylvia Frühwirth-Schnatter. Data augmentation and dynamic linear models. *Journal of Time Series Analysis*, 15:183–202, 1994.

[17] A Gelman and DB Rubin. Inference from iterative simulation using multiple sequences. *Statistical Science*, 7:457–511, 1992.

[18] Donald Geman and Stuart Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6:721–741, 1984.

[19] John Geweke. The dynamic factor analysis of economic time series. In Dennis J. Aigner and Arthur S. Goldberger, editors, *Latent Variables in Socio-Economic Models*. North-Holland, 1977.

[20] Rao K. Kadiyala and Sune Karlsson. Numerical methods for estimation and inference in Bayesian VAR-models. *Journal of Applied Econometrics*, 12:99–132, 1997.

[21] Chang-Jin Kim and Charles R. Nelson. *State-Space Models with Regime-Switching: Classical and Gibbs-Sampling Approaches with Applications*. MIT Press, Cambridge, Massachusetts, 1999.

[22] Eric Leeper, Christopher A. Sims, and Tao Zha. What does monetary policy do? *Brookings Papers on Economic Activity, Brookings Papers on Economic Activity, Economic Studies Program*, 27(2):1–78, 1996.

[23] Thomas J. Sargent and Christopher A. Sims. *Business cycle modeling without pretending to have too much a priori economic theory*. Federal Reserve Bank of Minneapolis, 1977.

[24] Christopher. Sims. The role of interest rate policy in the generation and propagation of business cycles: What has changed since the 30s? In Scott Shuh and Jeffrey Fuhrer, editors, *In Beyond Shocks: What Causes Business Cycles?* Federal Reserve Bank of Boston, 1999.

[25] Christopher A. Sims. Macroeconomics and reality. *Econometrica*, 48(1):1–48, 1980.

[26] Christopher A. Sims. Interpreting the macroeconomic time series facts : The effects of monetary policy. *European Economic Review*, 36(5):975–1000, 1992.

[27] Christopher A. Sims and Tao Zha. Were there regime switches in u.s. monetary policy? *American Economic Review*, 96(1):54–81, 2006.

[28] James H. Stock and Mark W. Watson. Implications of dynamic factor models for var analysis. *mimeo, Princeton University*, 1:1–100, 2005.

[29] James H. Stock and Mark W.Watson. Understanding changes in international business cycle dynamics. *NBER Working Paper 9859*, 2003.

[30] Harald Uhlig. What macroeconomists should know about unit roots: A bayesian perspective. *Econometric Theory*, 10:645–671, 1994.

[31] Harald Uhlig. What are the effects of a shock to monetary policy? results from an agnostic identification procedure. *Journal of Monetary Economics*, 52:381–419, 2005.

ERKLÄRUNG AN EIDES STATT

Ich erkläre hiermit an Eides Statt, daß die vorliegende Promotion selbständig und ohne die Hilfe Dritter verfaßt wurde und auch in Teilen keine Kopie anderer Arbeiten darstellt. Die benutzten Hilfsmittel sowie die Literatur sind als solche kenntlich gemacht und vollständig angegeben. Die vorliegende Promotion wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt.

Berlin, den 22.04.2009 Pooyan Amir Ahmadi